Universidade Nova de Lisboa
Instituto de Higiene e Medicina Tropical

Identification of proteases as diagnostic and drug targets in bovine babesiosis

Tiago Miguel Lopes Martins

Lisboa, 2009

Dissertation presented to obtain a PhD degree in Biomedical Sciences, speciality Parasitology, by the Universidade Nova de Lisboa, Instituto de Higiene e Medicina Tropical

The research described in this thesis was supervised by Doctor Ana Domingos and co-supervised by Doctor Vírgilio Estólio do Rosário. The work was performed at the Unidade de Tecnologias e Anticorpos Monoclonais, (formerly Instituto Nacional de Engenharia, Tecnologia e Inovação), Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Campus do Lumiar, Lisboa, Portugal; Centro de Malária e Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Rua da Junqueira, Lisboa, Portugal; and Faculdade de Veterinária, Universidade Eduardo Mondlane, Maputo, Mozambique.

Acknowledgments

The following thesis benefited from the insights and guidance of several people, to whom I am grateful.

I would like to express my gratitude to my supervisor, Dr. Ana Domingos, whose guidance, support, and patience, added considerably to my experience over the past years from a graduate student until the present time.

I am grateful to my co-supervisor, Professor Dr. Virgílio E. do Rosário for accepting me as student and for his unique guidance, planning, and whom I respect immensely.

I would like to thank the other member of my scientific committee, Dr. Sónia Centeno-Lima for the assistance she provided at all levels and especially in the initial development of this study.

To Dr. Luís Neves from the Faculty of Veterinary (Maputo, Mozambique), that received me in his laboratory, and for his endless support, advice and scientific enthusiasm that is an inspiration to whoever loves science.

To Dr. Varda Shkap from Kimron Veterinary Institute (Israel), for her support in this research project and her desire to contribute to a better knowledge on babesiosis.

To Dr. Carlos Novo, director of UTPAM, for the support, defense of his investigation group and constant instigation for us to acquire and improve our scientific abilities.

To Susana Teixeira for her work and dedication in the very beginning of the babesial adventure, when this research project was no more than a destiny.

To my present and former colleagues in UTPAM, Bé, Rubina, Raquel, Sofia, Fernando, Ana Armada, Ângela, Cristina, Cristina Baptista, Pedro, Albertina, Sandra; thanks for all the help, incentive, suggestions, patience, and above all friendship.

To my colleagues in Biotechnology Center of Faculty of Veterinary (Maputo, Mozambique) Olívia, Titus, and all the people that I have met there; that welcomed me to a second homeland, and for all the help, and friendship.

To my Mozambican "father", Ibraímo, that received me in his home as a son, and for his friendship.

À Lucia pelo amor e apoio constante que sempre me ilumina quando mais preciso.

Aos meus pais, familia e amigos que me ajudaram a crescer com valores, responsabilidade, amor e alegria.

Esta tese é dedicada ao irmão de um "cientista maluco", que sempre me incentivou a prosseguir este caminho depois de me ter guiado para um curso que adorei.

To the financial support provided by Fundação para a Ciência e Tecnologia (FCT) with a PhD fellowship grant (SFRH/BD/19059/2004) under POCI 2010 and the European Social Fund.

Ciência.Inovação 2010 Programa Operacional Ciência e Inovação 2010
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

GOVERNMENT OF THE PORTUGUESE REPUBLIC

EUROPEAN UNION
European Social Fund

Abstract

Bovine babesiosis is a tick–borne disease with significant morbidity and mortality, and the economic losses associated to this disease can be considerable. Control measures of bovine babesiosis include the eradication or reduction of ticks, good diagnosis, use of vaccination and correct treatment. The aim of this work was to contribute to a better diagnosis of the infection aiming at the improvement of some of the control measures, as well as to identify and characterize protease genes for the development of a diagnostic method and studied as potential drug targets.

For this study, collaborative work was carried out in Mozambique, from where blood samples from naturally infected cattle were collected in five farms located in the Maputo province, south of the country. A new molecular detection PCR method was then developed and tested using genomic DNA and random field samples collected from one farm. Primers were designed based in the babesial aspartic protease putative babesipsin gene identified in the genomes of *Babesia bigemina* and *B. bovis*. The new seminested hot-start PCR method was developed using the combination of 30 bp long primers and a hot start polymerase that theoretically allows the use of annealing temperatures above the melting temperatures of the primers and prevents the formation of unspecific amplifications and therefore increases the specificity.

The new seminested hot-start PCR method was assayed using 117 field samples in parallel with the widely used nested PCR method. The babesipsin seminested hot-start PCR was in this study more sensitive than the nested PCR. With the seminested hot-start PCR, 90% of the samples were positive for *B. bigemina* and 82% were positive for *B. bovis*. The results suggested that bovine babesiosis is common and endemic in Mozambique and that the disease was in an endemically stable situation.

The status of bovine babesiosis in Mozambique was then further studied, by testing random field samples from four more farms using the seminested hot-start PCR. All the samples from the five farms were also analysed using the reverse line blot (RLB) assay, and the results were compared with the data obtained by the seminested hot-start PCR. The detection of *Babesia* spp. differed considerably

between methods and locations. Using the seminested hot-start PCR, detection of *B. bigemina* between farms varied from 30% to 89% with an overall detection of 61%, and detection of *B. bovis* ranged between 27% and 83% with an overall frequency of 53%. Using the RLB assay *B. bigemina* was not detected and detection of *B. bovis* ranged between 0% and 17% with a total frequency of 5.1%. Analysis of new sequences of the 18S rRNA gene revealed that the current *B. bigemina* RLB probe is unspecific for the detection of all the identified isolates from Mozambique. The seminested hot-start PCR was therefore more sensitive than the RLB assay. Nevertheless, ten different species of the four genera *Anaplasma*, *Babesia*, *Ehrlichia* and *Theileria* were detected by the RLB assay, and this illustrates that multiple infections are widespread in Mozambique.

The results of this study show that bovine babesiosis is common in Maputo province, but there are some locations with low prevalence of infections and therefore the results suggest that this disease is not in an endemically stable situation in Maputo province. Further epidemiological studies are now needed to corroborate these findings.

Proteases have been shown to have essential roles in parasitic protozoa and are under study as promising drug targets. Some cysteine proteases of protozoan parasites are now recognized drug targets and specific inhibitors are in validation for chemotherapy of leishmaniasis, malaria, and trypanosomiasis. In this study our focus on the identification and characterization of proteases as drug targets was therefore in this class of proteases.

Cysteine protease putative genes were identified by sequence similarity search in the ongoing *B. bigemina* genome sequencing project database and were compared with the annotated genes from the complete bovine piroplasms genomes of *B. bovis*, *Theileria annulata*, and *T. parva*. Multiple genome alignments and sequence analysis were used to evaluate the molecular evolution events that occurred in the C1 family of cysteine proteases in these piroplasms of veterinary importance. There are five distinct groups of cysteine protease genes of C1 family in *B. bigemina* (5 genes), four groups in *B. bovis* (4 genes) and six groups in *Theileria* spp. (13 genes). Molecular evolution in *Theileria* occurred through the duplication of genes and sequence diversity. These considerable differences observed in the

cysteine protease family between *Babesia* and *Theileria* genera, may partially explain why *Babesia* cannot infect lymphocytes and *Theileria* infect first lymphocytes in the vertebrate host.

One of the identified cysteine proteases in the *B. bigemina* genome, babesipain-1, was expressed as a fusion protein with gluthatione S-transferase (GST) and the soluble protein was purified by affinity chromatography. The recombinant babesipain-1 showed activity against typical peptide substrates of cysteine proteases, and was inhibited by a general inhibitor of its class, but the low yield of the soluble purification prevented additional characterization.

Babesipain-1 was then purified from the insoluble fraction, and the denatured protein was refolded and activated to produce an active mature enzyme. Analysis of the activity of babesipain-1 revealed typical properties of a papain-family cysteine protease, including hydrolysis of typical papain-family peptide substrate, an acidic pH optimum (5.5-6.0), requirement for a reducing environment for maximum activity, and inhibition by standard cysteine protease inhibitors as E-64, leupeptin, ALLN and cystatin. The results suggest that babesipain-1 has a role in cytosol environment, since babesipain-1 retained high activity against peptide substrates at pH 7.5 (83% of maximum), an uncommon feature of cysteine proteases of parasitic protozoa.

Thus the results of this study demonstrate that bovine babesiosis is widespread in Maputo province in Mozambique, although the disease is not in an endemically stable situation. The results also suggest that cysteine proteases of *Babesia* spp. are promising drug targets for the development of an effective treatment of bovine babesiosis. In face of these results a plan for future work is associated. Some aspects and results from this work can be adjusted to other countries, including Portugal.

Resumo

A babesiose bovina é uma doença transmitida por carraças, que causa elevada morbidade e mortalidade, e provoca consideráveis perdas económicas devido aos esforços para controlar esta doença. As medidas de controlo da babesiose bovina incluem a erradicação ou redução de carraças, correcto diagnóstico, assim como tratamento e vacinação apropriados. Este trabalho tem como objectivo contribuir para um melhor diagnóstico da infecção e a consequente melhoria de algumas das medidas de controlo, bem como identificar e caracterizar genes de proteases utilizados para o desenvolvimento de um método de diagnóstico e estudados como potenciais alvos para fármacos.

Para este estudo, foi realizado um trabalho de colaboração em Moçambique, onde foram colhidas amostras de sangue de bovinos naturalmente infectados, em cinco explorações situadas na província de Maputo, no sul do país. Um novo método de detecção molecular por PCR foi desenvolvido e testado utilizando DNA genómico e amostras de campo aleatórias colhidas numa das explorações. Os iniciadores de PCR foram desenhados com base no gene putativo da protease aspártica babesipsina-1 identificado nos genomas de *Babesia bigemina* e de *B. bovis*. O novo *seminested hot-start* PCR foi desenvolvido utilizando a combinação de iniciadores longos de 30 pb de comprimento e uma *hot-start* polimerase, que permitem teoricamente a utilização de temperaturas de emparelhamento acima da temperatura de *melting*, impedindo assim a formação de amplificações não específicas, o que aumenta a especificidade do método.

O novo *seminested hot-start* PCR foi avaliado utilizando 117 amostras de campo, e em paralelo com um método amplamente utilizado, o nested PCR. O *seminested hot-start* PCR neste estudo foi mais sensível que o nested PCR. Com o *seminested hot-start* PCR, 90% das amostras foram positivas para *B. bigemina*, e 82% foram positivas para *B. bovis*. Os resultados sugeriram que a babesiose bovina é comum e endémica em Moçambique, e que a doença se encontra numa situação de estabilidade endémica.

O estudo do estado da babesiose bovina em Moçambique, foi então aprofundado, através da análise de amostras de campo aleatórias de mais quatro explorações

utilizando o *seminested hot-start* PCR. Todas as amostras das cinco explorações foram também analisadas utilizando o RLB, e os resultados deste método foram comparados com os dados obtidos pelo *seminested hot-start* PCR. A detecção de *Babesia* spp. diferiu significativamente entre os métodos utilizados e os locais de recolha. Com o *seminested hot-start PCR*, a detecção de *B. bigemina* nas várias explorações, variou entre 30% e 89%, com uma detecção total de 61%, e a detecção de *B. bovis* variou entre 27% e 83% com uma frequência global de 53%. Utilizando o RLB, não foi detectado *B. bigemina* e a detecção de *B. bovis* variou entre 0% e 17% com uma frequência total de 5,1%. A análise de novas sequências do gene 18S rRNA, revelou que a actual sonda do RLB para *B. bigemina* não é adequada para a detecção de todos os isolados desta espécie identificados em Moçambique. O *seminested hot-start* PCR foi portanto mais sensível que o RLB. No entanto, dez espécies diferentes dos quatro Géneros *Anaplasma*, *Babesia*, *Ehrlichia* e *Theileria* foram detectadas pelo ensaio RLB, e isso demonstra que as infecções múltiplas são comuns em Moçambique.

Os resultados deste estudo mostram que a babesiose bovina é comum na província de Maputo, e também que existem alguns locais com baixa prevalência de infecções, e portanto, os resultados sugerem que esta doença não está numa situação de estabilidade endémica na província de Maputo. São agora necessários novos estudos epidemiológicos para confirmar estes resultados.

Tem sido demonstrado que as proteases têm papéis essenciais em parasitas protozoários e estão sob estudo como promissores alvos de fármacos. Algumas proteases cisteínicas de parasitas protozoários, são já reconhecidos alvos de fármacos, e encontram-se em validação inibidores específicos para a quimioterapia da leishmaniose, da malária e da tripanossomíase. Neste estudo, o nosso principal interesse na identificação e caracterização de proteases como alvos de fármacos foi portanto nesta classe de proteases.

Foram identificados no banco de dados do projecto em curso de sequenciação do genoma de *B. bigemina*, os genes putativos de proteases cisteínicas em pesquisas por similaridade de sequência, que posteriormente foram comparados com os genes anotados nos genomas completos das espécies de piroplasmas bovinos *B. bovis*, *Theileria annulata* e *T. parva*. Para avaliar os eventos da evolução molecular

que ocorreram na família C1 de proteases cisteínicas, foram feitos alinhamentos múltiplos entre os genomas e análises das sequências obtidas destes piroplasmas de importância veterinária. Existem assim, cinco grupos distintos de genes de proteases cisteínicas da família C1 em *B. bigemina* (5 genes), quatro grupos em *B. bovis* (4 genes) e seis grupos em *Theileria* spp. (13 genes). No Género *Theileria* a evolução molecular ocorreu através da duplicação de genes e da diversificação da sequência das proteínas codificadas por estes genes. Estas importantes diferenças observadas entre os Géneros *Babesia* e *Theileria* na família das proteases cisteínicas, podem parcialmente explicar os diferentes mecanismos de infecção destas espécies, em que parasitas *Babesia* não invadem linfócitos e parasitas *Theileria* invadem primeiro os linfócitos no hospedeiro vertebrado.

A babesipaína-1, uma das proteases cisteínicas identificadas no genoma de *B. bigemina*, foi expressa como uma proteína de fusão com a glutationa S-transferase (GST) e a respectiva fracção solúvel foi purificada por cromatografia de afinidade. A babesipaína-1 recombinante apresentou actividade contra certos péptidos que são substratos típicos de proteases cisteínicas, e foi inibida por um inibidor geral da sua classe, mas o baixo rendimento da purificação da fracção solúvel impediu a sua caracterização adicional.

A babesipaína-1 foi então purificada a partir da fracção insolúvel, e a proteína desnaturada foi re-enrolada e activada para produzir uma enzima activa. A análise da actividade da babesipaína-1 revelou propriedades típicas de uma protease cisteínica da família da papaína, incluindo a hidrólise de alguns péptidos que são substratos típicos da família da papaína, um pH ácido óptimo (5.5-6.0), o requisito de um ambiente redutor para ter actividade máxima, e a inibição por inibidores de proteases cisteínicas como o E-64, a leupeptina, o ALLN e a cistatina. Os resultados sugerem que a babesipaína-1 tem um papel no citosol, já que a babesipaína-1 manteve elevada actividade contra substratos a pH 7,5 (83% do máximo), uma característica incomum das proteases cisteínicas de parasitas protozoários.

Assim, os resultados deste estudo demonstram que a babesiose bovina é uma infecção comum na província de Maputo em Moçambique, embora a doença não esteja numa situação de estabilidade endémica. Os resultados também sugerem que as proteases cisteínicas de *Babesia* spp. são alvos promissores para fármacos e

consequentemente para o desenvolvimento de um tratamento eficaz para a babesiose bovina. Na sequência destes resultados foi associado um plano de trabalho futuro. Alguns pormenores e resultados deste trabalho podem ser transferidos para outros países, inclusivamente Portugal.

Abbreviations

Abs	Absorbance
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride
AIDS	Acquired immunodeficiency syndrome
ALLN	N-acetyl-leucineleucine-norleucinal
AMC	7-amino-4-methyl coumarin
bp	Base pair
CP	Cysteine protease
DCI	3,4-dichloroisocoumarin
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
E-64	N-(trans-epoxysuccinyl)-l-leucine 4 guanidinobutylamide
E-64d	(2S,3S)-trans-epoxysuccinyl-L-leucylamido-3-methylbutane ethyl ester
EDTA	Ethylene diamine tetra acetic acid
ELISA	Enzyme-linked immunosorbent assay
EST	Expressed sequence tag
G+C	Guanine-cytosine content
GSH	Glutathione
GSSG	Glutathione disulfide
GST	Glutathione S-transferase
HIV	Human immunodeficiency virus
IFAT	Indirect immunofluorescence antibody test
IPTG	isopropyl-β-D-thiogalactopyranoside
LAMP	Loop-mediated isothermal amplification
LB-Lennox	Lysogeny broth, Lennox formulation
nPCR	Nested PCR
ORF	Open reading frame
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction

PMSF	phenylmethylsulfonyl fluoride
RLB	Reverse line blot hybridization
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SSU rRNA	small subunit ribosomal RNA
TBD	Tick borne disease
TBE	Tris/Borate/EDTA buffer
TLCK	Nα-Tosyl-L-lysine chloromethyl ketone hydrochloride
Tm	Melting temperature
TPCK	N-p-Tosyl-L-phenylalanine chloromethyl ketone
UV	Ultraviolet

Table of Contents

Index of figures

Index of Tables

Chapter 1

General Introduction

1.1. Introduction

Babesiosis caused by intraerythrocytic parasites of the Apicomplexan genus *Babesia*, is a common infection of vertebrate animals worldwide (Uilenberg, 2006). In 1893, Smith and Kilbourne discovered that *Babesia bigemina*, a causative agent of bovine babesiosis, was transmitted by the tick *Boophilus annulatus*. This revelation was a breakthrough in the history of parasitology, as it was the first proof that an arthropod was the vector of any disease agent (Uilenberg, 2006). Ticks from several genera are now known to be vectors and reservoirs of numerous *Babesia* spp. transmissible to reptiles, birds and mammals (Uilenberg, 2006). The major economic impact of babesiosis is on the cattle industry and the two most important species in cattle, *B. bigemina* and *B. bovis* threaten the health and safety of about 300 million cattle in tropical and subtropical regions of the world (Wright, 1990). The pathogenicity varies both between and within species but in many cases has a high rate of mortality in susceptible untreated animals (Bock et al., 2004). Many factors contribute to the emergence and re-emergence of babesiosis: parasite drug resistance, tick resistance to acaricides, genetic evolution, and the so-called global change (economic, social and environmental factors) (L'Hostis and Seegers, 2002). Climate change and increased climatic variability are particularly likely to affect vector borne diseases at the global scale in the near future. Control measures of bovine babesiosis include the eradication or reduction of ticks, better diagnosis, vaccination and treatment of clinical cases. The ongoing genome sequencing projects of *Babesia* spp., and more importantly, the vast knowledge acquired with the study of the closely related malaria pathogen may boost the study of *Babesia* diversity and drug resistance and the development of specific fast acting drugs, recombinant vaccines, and of more specific and sensitive diagnostic methods currently required for the control of babesiosis. This tick borne disease is also gaining increasing interest as an emerging zoonosis in humans, as the incidence of babesiosis has been rising steadily both in the U.S.A. and Europe (Homer et al., 2000; Hildebrandt et al., 2008). Human babesiosis may have previously been overlooked in many parts of the world due to its rarity or a lack of medical awareness and microbiological detection methods (Hildebrandt et al., 2008). On the other hand, the number of patients with risk factors is currently

increasing (e.g. immunocompromised individuals and organ transplant recipients) and asymptomatic but chronically infected blood donors may be a source of transfusion-transmitted babesiosis (Hildebrandt et al., 2008). New and old *Babesia* pathogens of veterinary and medical importance continue to emerge around the world and the substantial health impact of babesiosis on livestock and man is continuing.

1.2. The Apicomplexa

The eukaryotic phylum Apicomplexa comprises more than 5000 species of endoparasitic protozoa, including the *Plasmodium* parasites responsible for malaria, a major disease of poverty (Cavalier-Smith, 1993; Roos, 2005). Apicomplexan parasites infect virtually all animals (Levine, 1985), and while some parasite life cycles are relatively simple, involving only a single host, others require sexual recombination in a vector species for transmission. Sexual recombination is relevant in the production of diverse parasites within a species and this subject area is important in understanding drug resistance or immune mechanisms of defense. Some parasites are specialists, restricted to particular species and tissues, whereas others are generalists. For example, *Plasmodium falciparum*, which causes the most lethal form of malaria, infects only Hominids, and is transmitted only by anopheline mosquitoes. In contrast, *Toxoplasma gondii* can infect almost any tissue of warm-blooded animals. In the complex life cycle stages of Apicomplexan parasites, extracellular forms in the host include an apical complex that gives the phylum its name, including secretory organelles associated with host cell attachment, invasion, and establishment of an intracellular parasitophorous vacuole (Sibley, 2004). Different stages of development are also relevant for the study of drug targets and their selection.

1.3. The Genus *Babesia*

It was at the end of the 19th century that the Romanian biologist Victor Babeş discovered micro-organisms in erythrocytes of cattle in Rumania and associated them with bovine hemoglobinuria or red water fever (Babes, 1888). Currently it is know that members of the genus *Babesia* are one of the most ubiquitous and

widespread blood parasites in the world, second only to the trypanosomes (Homer et al., 2000). The genus *Babesia* belongs to the phylum Apicomplexa, class Sporozoasida, order Eucoccidiorida, suborder Piroplasmorina and family Babesiidae (Levine, 1971, 1985; Allsopp et al., 1994). The *Babesia* species *sensu stricto*, or true *Babesia*, are characterized by transovarial transmission in the vector tick and the limitation of infecting only erythrocytes in the host (Uilenberg, 2006). *Babesia* organisms are protozoan parasites that are all transmitted by ticks to their specific vertebrate host. Completion of a life cycle and therefore the maintenance of *Babesia* are completely dependent on both the tick and the vertebrate host (Mehlhorn and Shein, 1984) and therefore, the distribution of all the different *Babesia* species is governed by the geographical distribution of the tick vectors that transmit them. The *Babesia* spp. known to infect cattle, and their vectors, are listed in Table 1.1.

Babesia organisms can be visualized in a Giemsa stained infected blood smear. They are pyriform (pear shaped) or seen as round or amoeboid forms (Levine, 1985). Depending on the *Babesia* species, a host erythrocyte may be parasitized by single, paired, or multiple organisms. The size of the organisms varies depending on the species and are therefore classified as either small (1.0-2.5 µm; *B. bovis*, *B. gibsoni*, *B. ovis* and *B. divergens*) or large *Babesia* (2.5-5.0 µm; *B. bigemina*, *B. caballi* and *B. canis*), accordingly to the size of the trophozoites (Levine, 1985). Electron microscopy of extracellular forms reveals the presence of an apical complex with micronemes, rhoptries, apicoplast and other organelles (Kawai et al., 1986; Kawai et al., 1999) characteristic of the phylum Apicomplexa. *Babesia* extracellular infective forms, the sporozoites, move by gliding or body flexion, and lack flagella or pseudopodia, characteristic of some Apicomplexans (Levine, 1985).

Table 1.1. Known *Babesia* species of cattle, their ixodid tick vector genus and geographical distribution.

Species	Domestic host(s)	Vector genus	Distribution
B. beliceri [a]	Cattle	*Hyalomma*	Russia
B. bigemina	Cattle, buffalo	*Boophilus, Rhipicephalus*	Africa, America, Asia, Australia, Europe
B. bovis	Cattle, buffalo	*Boophilus, Rhipicephalus*	Africa, America, Asia, Australia, Europe
B. divergens	Cattle, reindeer	*Ixodes*	Europe
B. jakimovi [a]	Cattle, reindeer	*Ixodes?*	Siberia
B. major	Cattle	*Haemaphysalis*	Europe
B. occultans	Cattle	*Hyalomma*	Africa, Asia
B. ovata	Cattle	*Haemaphysalis*	Asia

[a] These species are not recognized by all scientists and some may not be valid. Adapted from Uilenberg (2006).

1.3.1. Life Cycle

The complex life cycle of *Babesia* spp. takes place in two hosts and sexual and asexual reproduction proceeds through at least three stages:

(i) gamogony – sexual development with formation and fusion of gametes inside the tick gut,
(ii) sporogony - asexual reproduction in tick salivary glands,
(iii) merogony - asexual reproduction in the vertebrate host.

Babesia spp. are not characterized by a life cycle that is specific for the genus (Mackenstedt et al., 1995). Therefore, we describe here the life cycle of *B. bigemina* (Fig. 1.1), which is similar to the life cycle of *B. bovis*.

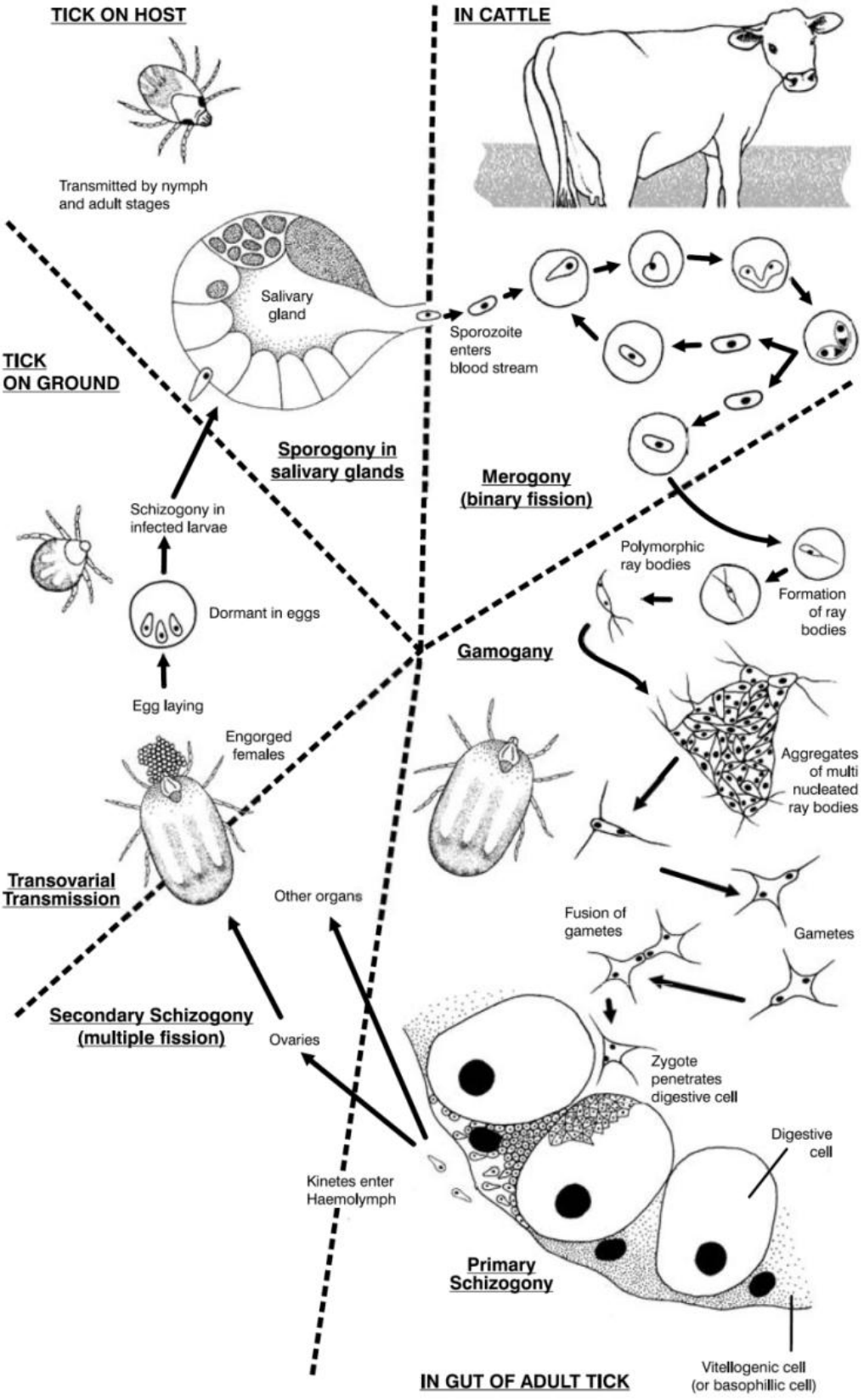

Figure 1.1. The life cycle of *Babesia bigemina* in cattle and the ixodid tick vector *Rhipicephalus* (*Boophilus*) *microplus*. Adapted from Bock et al. (2004).

1.3.1.1. Events in the vertebrate host

Once in the vertebrate, the transmitted bovine *Babesia* spp. sporozoites do not parasitise any vertebrate host cell other than erythrocytes (Mackenstedt et al., 1995). Each sporozoite penetrates the cell membrane of an erythrocyte with the aid of a specialised apical complex, forming a parasitophorous vacuole (Sibley, 2004). The vacuole membrane gradually disintegrates, and the parasite is left with the defining piroplasm feature of a single membrane, in contrast to *Plasmodium* species, which invade by a similar mechanism but retain the host membrane in addition to its own. Once inside the erythrocyte, it transforms into a trophozoite from which two merozoites develop by a process of merogony (Levine, 1985). The merozoites lyse the cell and go on to infect additional erythrocytes. Rapid reproduction destroys the host cell and leads to hemoglobinuria in the host. Some trophozoites develop into a diploid ovoid type of merozoite, called a gamont precursor. These gamont precursors do not develop further until they are taken up by the tick in the blood meal (Mackenstedt et al., 1995).

1.3.1.2. Events in the tick

Environment changes in the passage from host blood to the midgut of the tick vector stimulate the development of two populations of gamonts called ray bodies (Gough et al., 1998). The ray bodies undergo further multiplication within the consumed erythrocyte and form large aggregations of multinucleated ray bodies. After the consumed erythrocyte is digested, and once gametogenesis is complete, single-nucleated and haploid gametes (Mackenstedt et al., 1995) emerge from the aggregates and then fuse together in pairs to form a zygote (Gough et al., 1998; Mosqueda et al., 2004). The zygote selectively infects the digestive cell of the tick gut where they probably multiply and then the basophilic cells where further multiplication occurs. At some stage in development in the gut the zygote undergoes one-step meiosis to form a haploid zygote (Mackenstedt et al., 1995). In the gut cells, schizogony occurs with the formation of polyploid kinetes, and these motile club-shaped kinetes then escape into the haemolymph and infect a variety of cell types and tissues, including the oocytes where successive cycles of secondary schizogony take place (Mackenstedt et al., 1995). Consequently,

transovarial transmission occurs with further development taking place in the larval stage. This is an important life cycle adaptation as the *Boophilus* vectors are one-host ticks. Kinetes enter the salivary glands and are transformed into multinucleated stages (sporogony) and these then break up to form sporozoites (Mackenstedt et al., 1995). In all species, sporozoite development usually begins when the infected tick attaches to the vertebrate host. In *B. bigemina*, infective sporozoites take about 9 days to appear and therefore transmission only occur in the nymphal and adult stages of the tick (Hoyte, 1961; Potgieter and Els, 1977). In the case of *B. bovis*, the formation of infective sporozoites usually occurs within 2 to 3 days of larval tick attachment (Riek, 1966), and transmission occurs only in this stage as infective sporozoites do not persist beyond the larval stage (Mahoney and Mirre, 1979).

1.3.2. Epidemiology

Different breeds of cattle are known to differ in susceptibility to infection and manifestation of clinical signs of babesiosis. *Bos taurus* breeds are more susceptible to both *B. bigemina* and *B. bovis* than *Bos indicus* cattle (Bock et al., 1997). Age resistance was also demonstrated in calves from non immune mothers and this is related to passively acquired resistance from colostrum and innate immunity. Calves exposed to babesiosis during the first 6 to 9 months rarely show clinical symptoms and develop a solid long-lasting immunity (de Waal and Combrink, 2006), and if the infection recurs repeatedly the immunity is permanent. Under conditions of endemic instability, some animals will not be challenged for some periods after birth and may therefore develop severe, life threatening disease if they are exposed later in life (Bock et al., 2004). If at least 75% of calves are exposed to *B. bovis* infection by 6 to 9 months of age the disease incidence would be very low and a state of natural endemic stability would exist (Mahoney, 1974). One infected tick is sufficient to transmit *B. bovis*, but tick infection rates can be low and the rate of transmission to cattle depending on the breed is therefore usually slow. The use of tick control with acaricides unbalances the natural conditions, and numbers of infected ticks may fall below those required to maintain endemic stability. In endemic areas the animals clinically affected are

mostly susceptible cattle introduced for breeding purposes, for slaughter, or in transit. Severe clinical cases which occur in these cattle are caused by exposure to stresses such as parturition and starvation (Bock et al., 2004). Animals that are aged, splenectomized or immunocompromised are more susceptible to babesiosis than young and healthy animals (Levine, 1985).

1.3.3. Diagnosis and identification of parasites

Diagnosis of clinical cases of babesiosis is most frequently made by examination of blood smears stained with Giemsa or acridine orange (Böse et al., 1995; Trees, 1974). Blood films are prepared from capillary blood, as blood of the general circulation may contain up to 20 times fewer *B. bovis* parasites due to the sequestration of infected erythrocytes in the capillaries of the brain and other organs (Böse et al., 1995). In *B. bigemina* infections, parasitized cells are evenly distributed throughout the blood circulation. Thick blood films are 10 times more sensitive and are therefore very useful for the detection of low level *B. bovis* infections (Böse et al. 1995). These films differ from thin ones in that the blood is not spread over a large area and is not fixed before staining, thus allowing lyses of the red blood cells and concentration of the parasites (Böse et al. 1995). These techniques are inexpensive and reasonably portable. However, the accuracy of diagnosis relies on the training and skill of the microscopist. Low parasitaemias and the presence of different and morphological similar parasites (e.g. other *Babesia* spp. and also *Theileria* spp.) may adversely affect the proper classification of infections. Diagnosis of babesial infections can be further confirmed by serologic evaluation and polymerase chain reaction (PCR) based assays. These tests also have improved sensitivity to diagnose chronic or asymptomatic infections, for the purposes of research, epidemiological studies, export certification or where vaccine breakdowns are suspected.

Different serodiagnostic tests have been described. The indirect immunofluorescence antibody test (IFAT) is the most widely used, while the enzyme-linked immunosorbent assay (ELISA) is the test system which holds the greatest promise for the future (Böse et al. 1995). The advantage that this test has over IFAT is that interpretation of results is less subjective and it is easily

automated for large numbers of samples, although it requires more antigen and very defined assay procedures. Thus far, improvements to the ELISA have been limited as the quality of crude antigen preparations made from infected blood is generally poor. Potentially, most of the problems associated with crude antigens can be overcome by the production of recombinant antigens. Some ELISAs based on highly defined recombinant antigens have been described and show promise (Goff et al., 2006; Goff et al., 2008).

In cases that are difficult to diagnose by smear or serology, or when the detection of carrier animals with very low parasitaemias is required, the direct detection of parasites by PCR based assays is used. With the evolution of more sensitive PCR based techniques, several methods have been described including nested PCR (Figueroa et al., 1993), reverse line blot (RLB) hybridization (Gubbels et al., 1999), LAMP (Iseki et al., 2007) and real time PCR (Buling et al., 2007) for detection and differentiation of bovine babesiosis infections. Currently, none of these methods is used globally since some of them have not been validated to worldwide use, some require complicated post-PCR detection methods to further enhance the sensitivity or differentiation, others require special equipment, and also some may be prone to amplicon contamination issues.

1.3.4. Treatment and vaccination

Chemotherapy is generally effective against bovine babesiosis, with essentially the same drugs used for *B. bigemina* and *B. bovis*. Successful treatment depends on early diagnosis and the prompt administration of effective drugs. There is less likelihood of success if treatment is delayed until the animal has been weakened by fever and anaemia. Current treatments provide protection from clinical diseases but usually allow a sufficient level of infection (low level parasitaemias) for immunity to develop which is interesting in areas where babesiosis is endemic. Only a few babesiacides are now available commercially and diminazene aceturate and imidocarb dipropionate are the most widely used:

1) Diminazene – used intramuscularly at a dose of 3.5 mg/kg for treatment (de Vos, 1979).

2) Imidocarb - used subcutaneously at a dose of 1.2 mg/kg for treatment; is also the only chemoprophylactic, at a dose of 3 mg/kg, which will prevent clinical infection from *B. bovis* for 4 weeks and *B. bigemina* for at least 2 months (Taylor and McHardy, 1979).

At the high dose of 3 mg/kg, imidocarb may also completely eliminate *B. bovis* and *B. bigemina* leaving the animals susceptible to reinfection and for this reason reduced drug levels are sometimes indicated (Vial and Gorenflot, 2006, Bock et al., 2004), especially in endemic areas where the development of protective immunity is desirable. But, the use of reduced drug levels increases the risk of development of resistance against the drug when used extensively (Rodriguez and Trees, 1996). When a treatment is carried out on infected animals, and at a later stage parasites are detected again in the blood stream, several hypotheses can be made, as:

a) parasites under treatment were resistant to the administered drugs and therefore we are in presence of recrudescent parasite resistance;
b) a new infection has taken place, at a time where the drugs were not any longer in the blood stream and had no effect on the new parasites;
c) there has been no proper metabolization of the drug and therefore parasites were not totally affected by the treatment.

In any case, these are examples related to curative treatment and not prophylaxis. Imidocarb has a 28-day withholding period for meat consumption and restrictions for lactating cattle (the Chemical Safety Information from Intergovernmental Organizations, http://www.inchem.org/). Imidocarb at either dose can interfere with the development of immunity following live vaccination (de Vos et al., 1986). Treatment with long-acting oxytetracycline following vaccination significantly reduces parasitaemia and red blood cell destruction without inhibiting the development of immunity (Pipano et al. 1987). Oxytetracyclines are not usually able to control virulent field infections. Supportive treatment for chemotherapy is sometimes desirable, particularly in valuable animals. Blood transfusions may be life-saving in very anaemic animals. Anti-inflammatory drugs, such as

phenylbutazone, help relieve the inflammatory processes that occur, particularly with *B. bovis* infections (Vial and Gorenflot, 2006).

In addition to chemotherapy, life attenuated vaccines for cattle are used with over 95% efficacy from a single vaccination (Bock et al., 2004). Attenuated *Babesia* spp. strains used in vaccines are produced by different methods according to the species. Avirulent *B. bovis* parasites are produced by multiple passages of selected isolates in splenectomised calves, and avirulent *B. bigemina* are produced by splenectomising latently infected calves and using the ensuing relapse parasites to repeat the procedure in other animals (Bock et al., 2004). *In vitro* methods have also been used to grow attenuated organisms for preparation of live vaccines (Shkap and Pipano, 2001). While attenuated live vaccines lead to significant decreases in mortality, they do not prevent infection or full protection against milder symptoms of the disease (de Waal and Combrink, 2006). Live vaccines have been used routinely or experimentally in several countries (de Waal and Combrink, 2006; Bock et al., 2004; Shkap and Pipano, 2001). The observed decreasing efficiency of live vaccines in some regions, the complicated production methods, reduced shelf life and transportation issues, the major risks involved by using blood-derived live vaccines like transmission of contaminating pathogens and the reported heterogeneity between and within natural isolates urges the search for efficient recombinant vaccines (Bock et al., 2004). However the lack of understanding of immune mechanisms to primary and secondary infection and the fact that many protozoa have developed elaborate mechanisms such as antigen variation for evading host immunity remain obstacles to developing effective vaccines using recombinant technology (Jenkins, 2001).

1.4. Proteases as drug targets

Proteases catalyse the cleavage of amide linkages in macromolecular proteins and oligomeric peptides. Proteases have been identified in all biological systems from viruses to vertebrates. Once thought as only nonspecific digestive enzymes, proteases are now also known to be involved in substrate cleavages that result in changes of protein function. Proteases are involved in a wide range of biological processes such as DNA replication, cell signalling, immunity and apoptosis. From

genome sequencing data it is know that more than 550 proteases are defined in the human genome accounting for 2% of the expressed genes and similar percentages are present with little variance in all organisms (Puente et al., 2003; Wu et al., 2003). In addition, about 77 mutated human proteases have been identified to date which often contribute to hereditary diseases and, therefore, represent target opportunities (Quesada et al., 2009). Therefore, proteases are recognized as one of the largest potential drug target enzyme families. In addition, there are many more proteases found in viruses, bacteria, and parasites, which are also potential drug targets, due to their lower homology to their mammalian orthologs, and offer target opportunities to identify selective inhibitors that have minimal cross-reactivity with mammalian proteases (McKerrow et al., 2008). The inhibition of HIV-1 protease and human angiotensin-converting enzyme (ACE) involved in hypertension and congestive heart failure, demonstrated the potential of protease inhibitors in the treatment of infections and other diseases (Lüthi, 2002). Currently, several other proteases are in advance stages of study. Clinical development of protease inhibitors for diabetes, cancer, thrombosis, and osteoporosis is advancing. Phase III clinical trials are ongoing for the inhibitor Telaprevir (VX-950) of Hepatitis C virus NS3 protease and the inhibitor Rupintrivir (AG7088) of common cold human Rhinovirus 3C protease (Lüthi, 2002). The cysteine protease of the protozoan parasite *Trypanosoma cruzi*, cruzipain, is a validated target of effective inhibitors, and the drug candidate K777, is in late preclinical trials for Chagas disease (Doyle et al., 2007). Some HIV antiretroviral drugs have also been found to have anti-malarial and anti-giardiasis properties (Dunn et al., 2007; Martins et al., 2006). The widespread use of protease inhibitors as effective therapy for hypertension and AIDS, confirmed that proteases are valuable drug targets and current research will likely increase the number of available chemotherapeutic strategies based on the inhibition of proteases.

Drugs for treatment of infections must be very inexpensive for widespread use in resource poor regions. Many protease inhibitors can be made cheaply. The chemistry of protease inhibitors is so diverse that inexpensive synthetic schemes using simple starting materials can be developed or selected. The cost of HIV protease inhibitors is dependent on market considerations. HIV protease

inhibitors are being distributed to economically poor regions of the world at increasingly lower cost (WHO, Towards universal access, Scaling up priority HIV/AIDS interventions in the health sector, progress report 2008). Also, many viral or parasite infections will require only short courses of treatment, leading to much cheaper therapeutic regiments compared to the chronic treatment required for hypertension, diabetes or chronic viral infections (McKerrow et al., 2008).

1.4.1. Protease classification

Proteases (or peptidases) are divided into seven classes on the basis of the catalytic mechanism used during the hydrolytic process: (1) serine proteases, (2) threonine proteases, (3) cysteine proteases, (4) glutamic proteases, (5) metalloproteases, (6) aspartic proteases and (7) unknown. In the first three classes, an amino acid residue (serine, threonine or cysteine respectively) acts as the catalytic nucleophile that binds to the target peptide bond, whereas in the case of the glutamic, metallo and aspartic proteases a water molecule performs that role. Each of the main classes is subdivided into clans, which are further divided into families according to sequence homology (Rawlings et al., 2008).

1.4.2. Proteases of *Babesia* and other Apicomplexans

Since members of the Apicomplexa are obligate intracellular parasites, their survival depends on their ability to invade host cells, avoid degradation by host cell machinery and propagate intracellularly. Thus, in these organisms, proteases carry out tasks common to many eukaryotes as well as functions highly specific to the parasite life cycles. The Apicomplexa are named for the unique set of secretory organelles that are closely related with these functions. Data from diverse pathogens like *Plasmodium*, *Toxoplasma*, *Eimeria* and *Cryptosporidium* place the proteolytic processing by cysteine and serine proteases of parasite surface proteins, required for attachment and penetration, in the centre of parasite rupture and subsequent invasion of host cells (Binder and Kim, 2004; Koussis et al., 2009; Wickham et al., 2003; Smith et al., 2005; Fetterer et al., 2007). In *Plasmodium*, aspartic, cysteine and metallo- proteases are also important in the critical metabolic process of the digestion of haemoglobin in the acidic food

vacuole (Francis et al., 1997). Proteolytic enzymes play key roles in the life cycle of Apicomplexa protozoan parasites or the pathogenesis of diseases they produce, and therefore, proteases have been validated as drug targets in a number of parasitic infections including malaria, toxoplasmosis, and cryptosporidiosis (McKerrow et al., 2008).

An initial analysis of the *P. falciparum* genome revealed over 90 genes with homology to well characterized protease families (Wu et al., 2003) and *Toxoplasma gondii* is likely to encode as many protease genes (www.toxodb.org). The distribution by protease classes of the putative genes of *P. falciparum* (11% aspartic, 36% cysteine, 22% metallo, 17% serine, and 14% threonine) resembles those in other model organisms, and is noteworthy that cysteine proteases predominate as in many other parasitic protozoa (North et al., 1990). Therefore, it is not surprising that the early studies on proteases from Piroplasms were made in the identification of cysteine protease genes from *Theileria* spp. (Baylis et al., 1992; He et al., 2005; Holman et al., 2002; Nene et al., 1990; Sako et al., 1999). However, the identification of such enzymes or other proteases has not been explored in detail in *Babesia* species. Recently a serine protease from *B. divergens*, BdSUB-1, was identified and localized in the apical complex dense granules (spherical bodies), and anti-BdSUB-1 antibodies had an inhibitory effect on the invasion process (Montero et al., 2006). Serine proteases were also shown to be important in the invasion process and growth of *B. divergens* by inhibition of *in vitro* cultures with the serine class inhibitors TPCK and TLCK (Montero et al., 2007). It was also showed recently, that the lipophilic cysteine protease class inhibitors E-64d and ALLN reduced *in vitro* the invasion of erythrocytes as well as the growth of *B. bovis* (Okubo et al., 2007). Therefore, the possibility of developing selective inhibitors of key proteases of *Babesia* parasites into novel chemotherapeutic strategies can be explored with probable success.

1.5. Outline of the study

In this study, the identification of protease genes from bovine *Babesia* spp. was prosecuted in order to develop a new diagnosis method (Chapter 2 and 3) and to characterize potential drug targets for the future development of new therapeutic

strategies (Chapter 4 and 5). **Chapter 2** reports the development of a new molecular diagnostic method of bovine babesiosis denominated seminested hot-start PCR, using 117 field samples from Mozambican cattle, and based on the amplification of the aspartic protease babesipsin-1 gene. In **Chapter 3**, the current status of bovine babesiosis in Mozambique was further studied using the seminested hot-start PCR and the RLB assay to detect infections with *B. bigemina* and *B. bovis* in 477 samples from 5 different farms located in Maputo Province. **Chapter 4**, describes the identification of all the genes from the cysteine protease family C1 in the *B. bigemina* genome and the relationships with genes of the same family in the Piroplasms. For the first time, the activity of a cysteine protease from Piroplasms is reported. Recombinant babesipain-1 was expressed and purified as a soluble fusion protein with GST, and showed activity against peptide substrates. With this strategy, the amount of active enzyme produced was insufficient to proceed with further characterization. The production of sufficient amount was achieved with the purification of the insoluble GST-babesipain-1 fusion protein and the consequent refolding, activation and biochemical characterization as described in **Chapter 5**. In the final chapter, **Chapter 6**, the findings of the study described here are discussed and future potential studies are elaborated.

Chapter 2

Development of a new molecular diagnostic method for the detection of bovine babesiosis, its application in Mozambique

2.1. Introduction

Babesiosis is a tick borne disease (TBD) caused by parasites of the genus *Babesia*, with considerable worldwide economic, medical, and veterinary impact. In recent years, efforts have been made to rebuild the livestock population in Mozambique. Cattle have been imported mostly from the neighbouring countries Zimbabwe and South Africa. The success of this approach is being impaired by high mortality among the imported cattle, which is estimated at around 50% within the first year (Alfredo et al., 2005). Tick Borne Diseases, particularly babesiosis, anaplasmosis and cowdriosis are arguably the major causes of mortality. Although the distribution and prevalence of TBDs in Mozambique is mostly unknown, seroprevalence of *Babesia bovis* (39%) and *Anaplasma marginale* (63%) was recently reported for Tete province (Alfredo et al., 2005).

Bovine babesiosis in Africa, and particularly in Mozambique, is caused by *B. bovis* and *B. bigemina* (Uilenberg, 2006). There are reports of *B. occultans* in South Africa (Gray and De Vos, 1981) and possibly Nigeria (Dipeolu and Amoo, 1984), but the prevalence and distribution of this benign form of cattle babesiosis is unknown. On the other hand, the tick vector responsible for the transmission of *B. occultans*, *Hyalomma marginatum rufipes*, was identified in Mozambique, suggesting that *B. occultans* may therefore be present. *In vitro* cultivation of *B. occultans* was already accomplished (Van Niekerk and Zweygarth, 1996), but unfortunately there are no published sequences of *B. occultans* in the public databanks, and there is also no Polymerase Chain Reaction (PCR) detection method available. Nevertheless, molecular studies should consider the presence of this and possibly other unidentified species. Detection methods based on the 18S rRNA gene sequence or other genus conserved sequences, can originate false positives as these sequences are expected to be more conserved with the sequences from unsequenced and unidentified *Babesia* spp..

B. bovis was detected in Mozambique by serologic tests (Alfredo et al., 2005), but these methods are less sensitive and specific in the detection of the carrier state of animals and do not usually distinguish between past exposure and present infections. PCR based techniques constitute an alternative method for the direct detection of *Babesia* in carrier cattle. The carrier state occurs after acute or

primary infections, in which the animals are not clinically ill. Identification of carrier animals is important for the assessment of infection risk, given that they serve as reservoirs for infection of ticks and, ultimately, wider infection of the herd (Calder et al., 1996).

Several PCR based methods have been published that allow the detection of *B. bovis* and *B. bigemina*. More specifically, two methods have been used by various authors for the detection of *Babesia* from blood and ticks: the multiplex nested PCR (nPCR) for the detection of *B. bovis* and *B. bigemina* (Figueroa et al., 1993; Almeria et al., 2001; Gayo et al., 2003; Oliveira et al., 2005; Costa-Júnior et al., 2006) and the reverse line blot hybridization assay (RLB) (Gubbles et al., 1999; Georges et al., 2001; Brígido et al., 2004; Oura et al., 2004). Both methods appear to have similar sensitivities at 10^{-6} % parasitaemia (Gubbles et al., 1999; Costa-Júnior et al., 2006). However, it is interesting that the non-multiplex nPCR has an increased detection of 10^{-7} % parasitaemia (Oliveira-Sequeira et al., 2005).

The nPCR is a modification of the PCR procedure designed to increase the sensitivity. This modification consists of two sets of primers directed against the same target, used in two consecutive runs of PCR. The first set of primers is selected in the usual manner, whereas the second set of primers are localized internally or nested to the product of the first PCR run. A variation of the nPCR is the seminested PCR, where the second set of primers incorporates one of the primers of the first set (Winn Jr., 2006).

In this chapter, we report the successful development of a new method for the detection of *B. bovis* and *B. bigemina* in field samples and in face of the results the state of bovine babesiosis in Mozambique will be briefly addressed.

2.2. Materials and methods

2.2.1. Blood samples from cattle

A total of 117 blood samples were collected from cattle in the province of Maputo, Mozambique. The samples were collected in September near Umbelúzi in the Boane District, mainly from Friesian cattle. Approximately 2–4 ml of blood were collected from the coccygeal vein into ethylenediaminetetraacetic acid (EDTA)

buffered vacutainer tubes. Samples were kept at 4 °C while being transported to the laboratory at the Faculty of Veterinary of Maputo. The blood was stored at -20 °C until DNA extraction.

2.2.2. DNA extraction

DNA extraction was performed according to Centeno-Lima et al. (2003). Two hundred microlitres of EDTA buffered whole blood was added to 500 µl phosphate-buffered saline (PBS), vortexed for 10 s and then centrifuged at 16000 g for 5 min. The cells pellet was washed with PBS three more times or until the supernatant was clear. The pellet was then ressuspended with 100 µl of lysis buffer [50 mM KCl, 0.5% (v/v) Tween-20, 10 mM Tris-HCl (pH 8.0) and 10 µg of proteinase K added before use], incubated overnight in a water bath at 56 °C and heated for 10 min at 100 °C to inactivate proteinase K. Samples were stored at -20 °C.

2.2.3. PCR reactions

2.2.3.1. Primers

The primers used in hot-start PCR amplification (Table 2.1), are localized within the putative aspartic proteinase babesipsin genes from both *B. bovis* and *B. bigemina*. The babesipsin putative gene sequences were identified in the Sanger Institute databases: in the *B. bovis* EST Sequencing Project (de Vries et al., 2006) and in the *B. bigemina* genome project (http://www.sanger.ac.uk/Projects/).

The oligonucleotide primers were designed using the online GeneFisher program (Giegerich et al., 1996) with the following parameters: G+C content from 40 to 60%, melting temperature between 60 and 80 °C, and primer size between 27 and 31 bp (Table 2.1). The expected size using babesipsin primers in hot-start PCR is 614 bp for *B. bigemina* and 426 bp for *B. bovis*. The amplification products span 2 partial exons and 1 intron. The seminested hot-start PCR products have the expected size of 469 bp and 275 bp respectively for *B. bigemina* and *B. bovis*.

The primer sequences described by Figueroa et al. (1993) were used in nPCR; BoF/R and BilA/B the outer primers, BoFN/RN and BilAN/BN the inner or nested primers (Table 2.1).

Melting temperatures were calculated using the Wallace–Itakura rule (Wallace et al., 1979).

Table 2.1. Primers sequence and melting temperature (Tm).

Method and parasite	Name	Primer sequence (5´-3´)	Length (bp)	Tm (°C)
hot-start PCR *B. bigemina*	BigBAF1	GGGAGATAAAAATCGGCACGCCCCCGCAA	29	66
	BigBAR1	GAGGATCTATGCCTCCTAACATTATCCGTGA	31	62
	BigBAR2	GCATTCCGGGAACACTGCTCATTTCTGGGA	30	64
hot-start PCR *B. bovis*	BovBAF1	CCCGCTCTGGATACCGTAACCATAGGAGA	29	64
	BovBAR1	ATACTGAGGATCCACTCCGCCTAGCATCA	29	63
	BovBAR2	GCATTCCGGGTATGCTACTCATCTCTGGA	29	63
nPCR *B. bigemina*	BiIA	CATCTAATTTCTCTCCATACCCCTCC	26	58
	BiIB	CCTCGGCTTCAACTCTGATGCCAAAG	26	61
	BiIAN	CGCAAGCCCAGCACGCCCCGGTGC	24	69
	BiIBN	CCGACCTGGATAGGCTGTGTGATG	24	61
nPCR *B. bovis*	BoF	CACGAGGAAGGAACTACCGATGTTGA	26	60
	BoR	CCAAGGAGCTTCAACGTACGAGGTCA	26	61
	BoFN	TCAACAAGGTACTCTATATGGCTACC	26	56
	BoRN	CTACCGAGCAGAACCTTCTTCACCAT	26	60

2.2.3.2. Babesipsin hot-start PCR and seminested hot-start PCR

The babesipsin hot-start PCR reaction mixtures (20 µl) contained 16 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl (pH 8.8), 0.01% (v/v) Tween-20, 2.5 mM $MgCl_2$, 0.5 µM each primer F1 and R1, 200 µM each dNTP, 1 U of Superhot *Taq* DNA polymerase (Bioron GmbH), and 1 µl total DNA. Hot-start PCR was carried out in a PTC-200 MJ Research thermocycler for 40 cycles. Each cycle consisted of 20 s of denaturation at 95 °C (1 min for the first cycle), 30 s of annealing at 69 °C, and 45 s of extension at 72 °C. The same conditions of the first PCR were used in seminested hot-start PCR apart from using the inner reverse primers R2 instead of R1, 1 µl of the first PCR products as template and an annealing temperature of 69 °C for *B. bovis* and

71 °C for *B. bigemina*. Only the samples that were negative in the first hot-start PCR were submitted to seminested hot-start PCR. Twenty random PCR products were gel purified and sequenced in outsourcing at STAB Vida, Portugal.

2.2.3.3. nPCR

The first PCR using the primers described by Figueroa et al. (1993), was carried out using the same buffer as described for babesipsin hot-start PCR, 1 U DFS-*Taq* DNA polymerase (Bioron GmbH), 0.5 µM outer primers, 1 µl total DNA and the thermocycler program described by Oliveira et al. (2005): 35 cycles (1 min at 95 °C, 1 min at 60 °C for *B. bovis* and 64 °C for *B. bigemina*, 1 min 30 s at 72 °C) and a final extension step at 72 °C for 5 min. The same conditions of the first PCR were used in nPCR apart from using inner primers, 2 µl of the first PCR products as template and an annealing temperature of 65 °C for *B. bovis* and 70 °C for *B. bigemina*.

2.2.4. Analysis of PCR products

Ten microlitres of first PCR and 5 µl of second PCR products were separated by electrophoresis in 1.2% (w/v) agarose gel containing ethidium bromide in 0.5 x TBE buffer (44.5 mM Tris-HCl, 44.5 mM Boric Acid and 1 mM EDTA, pH 8.3). After electrophoresis, PCR products were visualized by transillumination with UV light.

2.2.5. Statistical analysis

The data obtained from DNA amplification (nPCR and seminested hot-start PCR) was used for comparisons among methods by Kappa coefficient analysis, a statistical measure of concordance for qualitative items, that takes into account the agreement occurring by chance.

2.3. Results

2.3.1. Hot-start PCR

The combinations using the babesipsin primers described in this study and genomic DNA (kindly provided by Dr. Varda Shkap from Kimron Veterinary Institute, Israel), allowed the amplification of single band products with the

expected size. There was no cross reaction between primers for one species and genomic DNA of the other species. And even if it occurs with other isolates, it is expected that the amplification products will have different sizes, since there is an intron between the locations of the primers that is 200 bp shorter for *B. bigemina*. Hot-start PCR using field samples also allowed the amplification of single band products with the expected size (see Fig. 2.1 and 2.2).

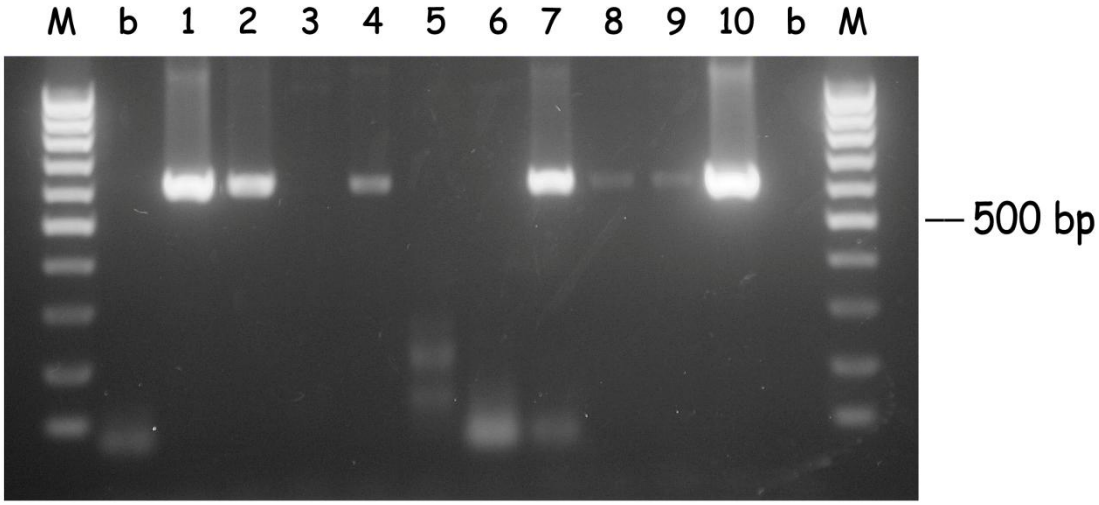

Figure 2.1. Hot-start PCR of *B. bigemina* DNA from random samples (1 to 10). DNA was subjected to hot-start PCR for the babesipsin 614 bp sequence amplification using the primers BigBAF1 and BigBAR1. Samples 3, 5 and 6 were negative and the remaining were positive. M: 100 bp DNA ladder marker; b: no DNA, negative control.

After the seminested hot-start PCR (see Fig. 2.3 and 2.4), 14 samples that were negative in the first PCR were positive for *B. bigemina*, and 15 were positive for *B. bovis*. Thus, the detection of *B. bigemina* and *B. bovis* after the seminested hot-start PCR increased and were respectively 90% and 80%. The mixed infections also increased to 86 samples (74%) and the number of negatives for both *Babesia* decreased to 2 (1.7%). Only one strong amplification with a different size was observed in the *B. bigemina* seminested hot-start PCR, and was considered as negative (Fig. 2.3, sample 2).

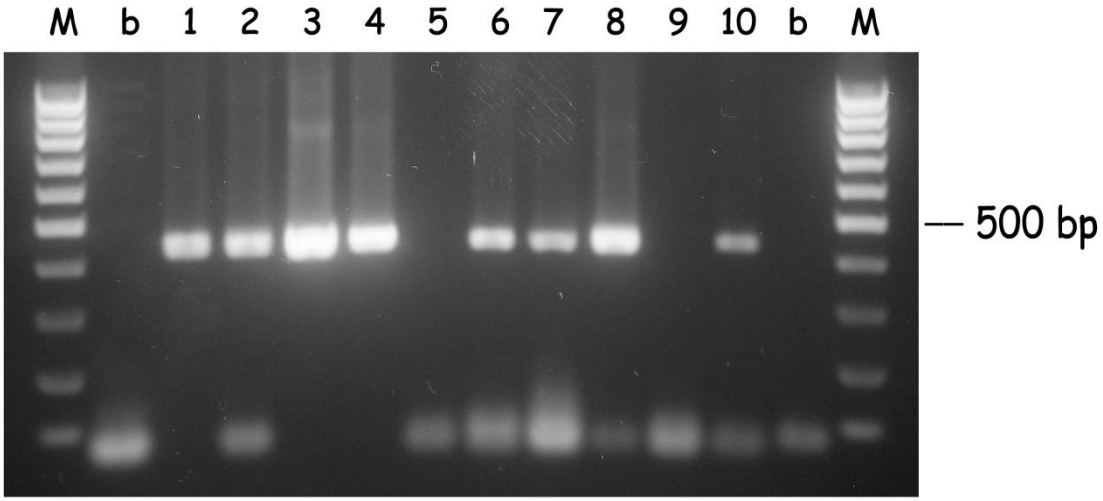

Figure 2.2. Hot-start PCR of *B. bovis* DNA from random samples (1 to 10). DNA was subjected to hot-start PCR for the babesipsin 426 bp sequence amplification using the primers BovBAF1 and BovBAR1. Samples 5 and 9 were negative and the remaining were positive. M: 100 bp DNA ladder marker; b: no DNA, negative control.

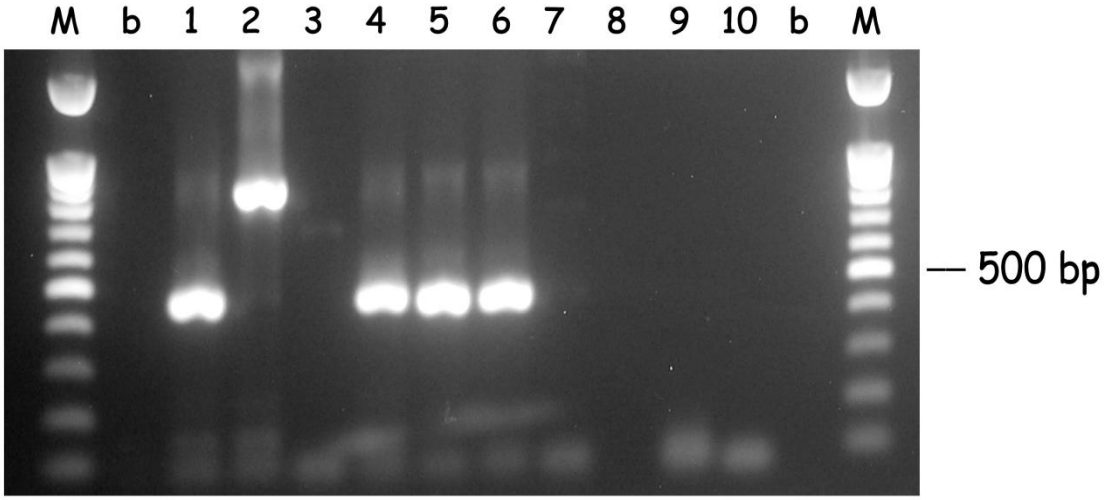

Figure 2.3. Seminested hot-start PCR of *B. bigemina* DNA from random samples (1 to 10). DNA was subjected to seminested hot-start PCR for the babesipsin 469 bp sequence amplification using the primers BigBAF1 and BigBAR2. Samples 2, 3 and 7 to 10 were negative and the remaining were positive. M: 100 bp DNA ladder marker; b: no DNA, negative control.

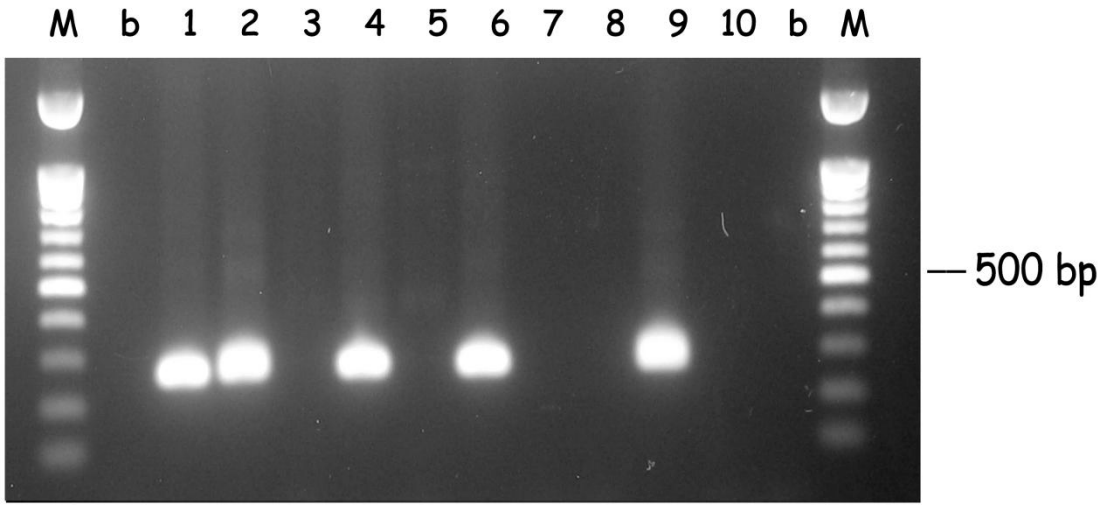

Figure 2.4. Seminested hot-start PCR of *B. bovis* DNA from random samples (1 to 10). DNA was subjected to seminested hot-start PCR for the babesipsin 275 bp sequence amplification using the primers BovBAF1 and BovBAR2. Samples 3, 5, 7, 8 and 10 were negative and the remaining were positive. M: 100bp DNA ladder marker; b: no DNA, negative control.

2.3.2. nPCR

All 117 samples were analyzed by nPCR for comparison with the hot-start PCR method. The nPCR reactions using the primers described by Figueroa et al. (1993), resulted in amplification from field samples of the desired products, but showed some unspecific background amplification when using *B. bigemina* primers (see Fig. 2.5). The unspecific amplification was in the same range as the desired product, making difficult the interpretation of results, mainly for positive samples. The unspecific amplification can be reduced using half the volume of template in the nPCR (1µl), but the detection decreased in less two positive samples, when all the samples were analysed with this modification. There was no amplification in the first PCR (not even using hot-start polymerase), most possibly due to low concentrations of parasite DNA template. In the second or nPCR, 90 samples were positive for *B. bigemina* (77%) and 91 were positive for *B. bovis* (78%).

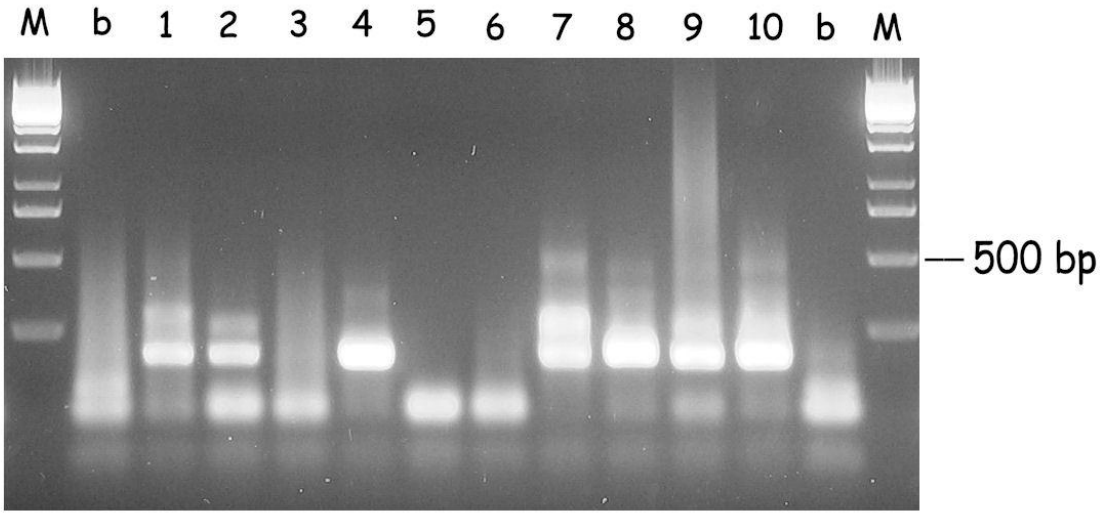

Figure 2.5. Nested PCR of *B. bigemina* DNA from random samples (1 to 10). DNA was subjected to nested PCR for the 169 bp sequence amplification using the primers BiIAN and BiIBN. Samples 3, 5, and 6 were negative and the remaining were positive. M: 1kbp DNA ladder marker; b: no DNA, negative control.

2.3.3. Statistical analysis

In comparison with the nPCR method, the seminested hot-start PCR allowed an increased detection for both *B. bigemina* and *B. bovis*. Some samples were only positive using the seminested hot-start PCR method, and some were only positive using the nPCR. Despite this, the observed mutual agreement between the two methods is relatively high: 0.829 in the detection of *B. bigemina* and 0.846 in the detection of *B. bovis*. The Kappa coefficient determined for the two detection methods of *B. bigemina* (0.412) was lower than for the detection methods of *B. bovis* (0.515).

2.4. Discussion

Babesiosis and other TBD´s may be responsible for severe economic losses in Mozambique, but little is known about the epidemiology of these diseases in this country. Here we report for the first time the direct detection of *B. bigemina* and *B. bovis* in field samples from Mozambique. Bovine babesiosis was detected by seminested hot-start PCR, in 90% of cattle with *B. bigemina* and in 82% with *B.*

bovis. It's also important to note that almost the entire herd was positive (98%) for *B. bigemina* and/or *B. bovis*. These results clearly indicate that this region of Mozambique is endemic for bovine babesiosis. Similar results were observed by Oliveira et al. (2005) in São Carlos, SP, Brazil, where they detected infection in cattle by *B. bigemina* (88%) and *B. bovis* (92%) using the nPCR. Although other studies should be conducted in other areas in Mozambique to confirm the endemic status of babesiosis in this country, efforts to rebuild the stock population should consider the babesiosis threat, especially when importing cattle from countries that adopt stricter control measures.

Table 2.2. Comparison of methods used for the detection of cattle infected with *B. bigemina* and/or *B. bovis* in Boane district, Mozambique.

Method	*B. bigemina*	*B. bovis*	Mixed infections	*No infection*
PCR	0 (0%)	0 (0%)	0 (0%)	0 (0%)
nPCR	90 (77%)	91 (78%)	73 (62%)	7 (6.0%)
hot-start PCR	90 (77%)	82 (70%)	61 (52%)	6 (5.1%)
seminested hot-start PCR	104 (90%)	97 (82%)	86 (74%)	2 (1.7%)

There are several reported methods in the literature describing the detection of *B. bigemina* and *B. bovis*, but only some have been tested with random field samples: the nPCR (Figueroa et al., 1993; Almeria et al., 2001; Gayo et al., 2003; Oliveira et al., 2005; Costa-Júnior et al., 2006), the RLB (Gubbles et al., 1999; Brígido et al., 2004; Oura et al., 2004) and more recently the LAMP assay (Iseki et al., 2007). All three methods have sensitivities around 10^{-6} % parasitaemia (Gubbles et al., 1999; Costa-Júnior et al., 2006; Iseki et al., 2007), although Iseki and co-workers (2007) reported a lower sensitivity for the nPCR (10^{-4} % to 10^{-5} % parasitaemia) in their comparative study with the LAMP assay.

In this study the seminested hot-start PCR was more sensitive than the nPCR (Table 2.2), although the kappa coefficients determined to compare the two assays indicate moderate agreement (0.412 and 0.515) and the observed mutual agreement between methods is relatively high (0.829 and 0.846). The disagreement between methods is not only the consequence of different detection

values. It is also the result of the observation that some samples were only positive using the seminested hot-start PCR method, and some were only positive using the nPCR. This error associated with each method can be expected when the samples have low parasitaemias, because the frequency of detection decreases with the decrease in parasitaemia (Calder et al., 1996). More significant, are the differences between methods in the first PCR (Table 2.2). The hot-start PCR improved considerably the detection in the first PCR, and this is a major advantage of this method. Unfortunately the primer pairs used didn't allow an improved detection between the first hot-start PCR and the nPCR. But the *B. bigemina* primer pair used in the first hot-start PCR showed the same sensitivity as the two step nPCR (Table 2.2). Other primer pairs could even increase the detection limit in the first hot-start PCR, especially if the target gene is changed to a multicopy gene instead of the apparent single copy babesipsin gene. With the release of the complete and annotated *B. bovis* and *B. bigemina* genomes, it will be possible to verify if there are multicopy genes suitable for this intra genus analysis. The 18s rRNA multicopy gene, is an alternative, but this gene is less variable and there is still no sequence information for *B. occultans*.

In the first PCR of the nPCR method, using BoF and BoR or BilA and BilB primers, there was no amplification, even when using hot-start polymerase. When using the babesipsin longer primers with higher melting temperatures (Table 2.1) and a non hot-start polymerase, there was also no amplification. Only the combination of a hot-start polymerase and long primers (29-31 bp) allowed the amplification from field samples with low concentrations of parasite DNA. This observation indicates that this combination increases the sensitivity of the PCR, probably by increasing the specificity. The use of hot-start polymerases and 30 bp primers with melting temperatures around 64 °C can eventually be used in the development of new detection methods for pathogens that are present in low numbers as in the case of the carrier state of *B. bovis* and *B. bigemina*.

In conclusion, our new method allowed the detection of *B. bigemina* and *B. bovis* in field samples from Mozambique and can now be used in further epidemiological studies to elucidate the endemic status of babesiosis in Mozambique.

Chapter 3

Bovine babesiosis and other cattle haemoparasites infections in Maputo province, Mozambique as diagnosed by molecular techniques

3.1. Introduction

Bovine babesiosis is a tick-borne disease (TBD) of cattle caused in Mozambique by the protozoan parasites *Babesia bigemina* and *B. bovis*. Bovine babesiosis is of great veterinary importance as this TBD cause severe economic losses throughout the world (Bock et al., 2004). In trypanosome free regions, TBDs are considered the main cause of morbidity and mortality of cattle in Mozambique, although the information on TBDs in the country requires an urgent update. Recently, two epidemiological studies on bovine babesiosis in Mozambique have been published: *Babesia bovis* (39%) was detected in a serological survey in Tete province (Alfredo et al., 2005); and, bovine babesiosis was detected in Maputo province using a new seminested hot-start PCR method (**Chapter 2**) that showed improved detection over a previous nested PCR method (Figueroa et al., 1993). 117 random field samples from one farm located in Maputo province were analysed by the seminested hot-start PCR and both *B. bigemina* and *B. bovis* were detected in over 80% of the samples (**Chapter 2**). These results indicated that the area of Maputo province, Mozambique, could be endemically stable for these species of bovine babesiosis. In the current chapter, the detection of bovine babesiosis by the seminested hot-start PCR and the reverse line blot (RLB) assay (Gubbels et al., 1999; Schouls et al., 1999; Bekker et al., 2002) in five farms geographically distributed in Maputo province is reported. The RLB assay is based on a reverse hybridization where the specific probes are immobilized to a solid support and the sample is denatured and subsequently allowed to hybridize with the probes. The RLB assay was also used to determine the presence of haemoparasites of the genera *Babesia*, *Theileria*, *Anaplasma* and *Ehrlichia*.

3.2. Materials and methods

3.2.1. Cattle

Blood samples were collected randomly from cattle in five different locations in the province of Maputo, Mozambique (Fig. 3.1): 117 samples in September near Umbelúzi, Boane district; 77 samples in July near Bela Vista, Matutuine District; 95 samples in October in Catuane, Matutuine District; 93 samples in October in Chobela, Magude District; and 95 samples in November in Namaacha, Namaacha District (Friesian cattle in Umbelúzi, and in the remaining farms Nguni and Nguni crosses). Blood was collected and processed as described elsewhere (**Chapter 2**).

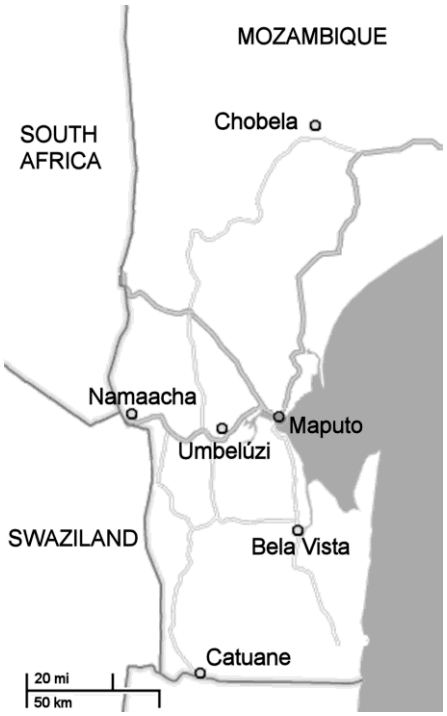

Figure 3.1. Map of the locations in Maputo province, Mozambique where samples were collected.

3.2.2. DNA extraction

DNA extraction was carried out as described previously (**Chapter 2**). Briefly, whole blood was washed with phosphate-buffered saline (PBS), and the remaining cells pellet were lysed in an appropriate buffer with Proteinase K, overnight at 56 °C.

3.2.3. Hot-start PCR and seminested hot-start PCR

Hot-start PCR was performed as described previously (**Chapter 2**) with minor changes. Hot-start PCR reaction mixtures (20 µl) contained 5 mM $(NH_4)_2SO_4$, 15 mM Tris-HCl (pH 8.2), 30 mM KCl, 2.5 mM $MgCl_2$, 0.02% (w/v) bovine serum albumine, 0.5 µM each primer F1 and R1, 200 µM each dNTP, 0.025 U of DyNAzyme™ II Hot Start DNA polymerase (Finnzymes Oy, Espoo Finland), and 2 µl of DNA. Hot-start PCR was executed for 40 cycles: 30 s of denaturation at 94 °C (10

min in the first cycle to activate the hot start polymerase), 30 s of annealing at 68 °C, and 45 s of extension at 72 °C. Negative samples in the first hot-start PCR were submitted to seminested hot-start PCR, using the conditions of the first PCR with the inner reverse primers R2 in substitution of R1, and 1 µl of the first PCR products as template. PCR products were examined in ethidium bromide-stained agarose gels.

3.2.4. Reverse line blot assay

The PCR products that were later hybridized with the probes in the RLB membrane were amplified with two sets of primers. For the identification of *Babesia* and *Theileria* species a fragment of 460–540 bp of the 18S SSU rRNA gene spanning the V4 region was amplified with primers described by Gubbels et al. (1999). For the identification of *Anaplasma* and *Ehrlichia* species a PCR product of 492–498 bp fragment of the 16S rRNA gene spanning the V1 region was amplified using primers described by Bekker et al. (2002). Primers were obtained from Isogen Life Science (Maarsen, The Netherlands). PCR amplification was performed in a 25 µl reaction volume consisted of 2.5 U Taq DNA polymerase (Invitrogen, USA), 1x PCR buffer, 50 pmol of each primer, 200 µM each dNTP and 2.5 µl of the template DNA with an automated cycler (I-Cycler, Bio-Rad) for 50 cycles: 10 cycles of 94°C for 20 s, 67°C for 30 s, and 72°C for 30 s, with a 2°C decrease of the annealing temperature every second cycle (touchdown); and forty cycles of 94°C for 20 s, 57°C for 30 s, and 72°C for 30 s. The RLB membrane with 34 species specific amino-linked oligonucleotide probes from Isogen Life Science (Maarssen, The Netherlands) was hybridized with the PCR products according to Nijhof et al. (2005). Each membrane was used for a maximum of 10 times.

3.2.5. Cloning of 18S ribosomal RNA genes

The partial *Babesia* spp. 18S rRNA genes were amplified by PCR. A PCR product with the expected size of approximately 430 bp was obtained with primers specific for *Babesia* spp. (Hilpertshauser et al., 2006). A longer PCR product with approximately 790 bp and a shorter product with 350 bp were respectively amplified using forward primers 5-22F and 455-479F with the reverse primer 793-772R (Birkenheuer et al., 2003). PCR products were cloned into the plasmid

vector pCR2.1 (Invitrogen), and positive clones were sequenced at STAB Vida (Oeiras, Portugal). Eleven unique sequences obtained in this study were submitted to GenBank under accession nos. FJ869895 to FJ869905. Multiple nucleic sequence alignment between 18S rRNA genes was performed using ClustalW (Thompson et al., 1994), followed by manual editing in Bioedit (Hall, 1999) and phylogenetic relationships constructed and visualized by MEGA4 program (Tamura et al., 2007).

3.3. Results

3.3.1. Seminested hot-start PCR

Detection levels of *Babesia* species differed considerably in the five farms studied. *B. bigemina* detection by the seminested hot-start PCR varied between a minimum of 30% and a maximum of 89%, with an overall detection of 61%. *B. bovis* detection changed between 27% and 82%, with an overall frequency of 53%. *B. bovis* was only detected in higher levels than *B. bigemina* in the Namaacha farm. A high percentage of the herds from the five farms were infected with either parasite, and only 25% of the 477 samples were negative for infection with bovine babesiosis. Mixed infections were common, with 37% of the cattle infected by both *Babesia* species. Considering only the results obtained in this study (excluding Umbelúzi farm), the seminested improved the detection over the first hot-start PCR in 2% for *B. bigemina* and in 11% for *B. bovis*.

3.3.2. Reverse line blot assay

The RLB hybridization method allowed the identification of ten different species of four genera of haemoparasites present in the region of Maputo province (Table 3.1). *B. bovis* was detected in four of the five farms, but in much lower percentages compared with the results obtained by the seminested hot-start PCR. *B. bigemina* was not detected by the RLB method. In Chobela, 3% of the samples tested positive for *B. vogeli*, known for infections in domestic dogs. *Anaplasma bovis* and *Theileria taurotragi* were detected in only two farms, whereas the other detected haemoparasites were widely distributed. The most frequently detected species using the RLB were *T. velifera* and *T. mutans*, present in respectively 54% and 52% of the samples.

Table 3.1. Prevalence of cattle haemoparasites infections in Maputo province, Mozambique.

	District	Boane	Magude	Matutuine		Namaacha
	Location	Umbelúzi	Chobela	Bela Vista	Catuane	Namaacha
	Month	September	October	July	October	November
	Breed	Friesian	Nguni and Nguni crosses			
	nº samples	117	93	77	95	95
hot PCR	*B. bigemina*	76.9%[a]	67.8%	29.9%	53.7%	51.6%
	B. bovis	70.1%[a]	43.0%	22.1%	21.1%	54.7%
seminested hot PCR	*B. bigemina*	88.9%[a]	72.0%	29.9%	55.8%	54.7%
	B. bovis	82.9%[a]	54.8%	27.3%	32.6%	68.4%
	mixed	73.5%[a]	44.1%	10.4%	20.0%	38.9%
	no detection	1.7%[a]	17.2%	55.8%	31.6%	15.8%
RLB	*B. bigemina*	0.0%	0.0%	0.0%	0.0%	0.0%
	B. bovis	17%	2.2%	2.6%	0.0%	4.2%
	B. vogeli	0.0%	3.2%	0.0%	0.0%	0.0%
	T. mutans	42.3%	69.9%	71.8%	22.1%	55.8%
	T. velifera	70.5%	51.6%	70.5%	23.2%	51.6%
	T. taurotragi	3.8%	0.0%	0.0%	0.0%	1.1%
	T. spp. Sable	11.5%	26.9%	3.8%	0.0%	1.1%
	A. centrale	6.8%	10.8%	0.0%	1.1%	2.1%
	A. marginale	23.9%	10.8%	7.7%	0.0%	1,1%
	A. bovis	3.3%	1.1%	0.0%	0.0%	0.0%
	E. spp.	12.8%	14.0%	0.0%	1.1%	2.1%

a – from **Chapter 2**.

3.3.3. 18S ribosomal RNA gene sequences analysis

Six different sequences of the *B. bigemina* 18S rRNA gene were obtained from samples collected at Chobela farm (Chobela1 to 6 with GenBank accession nos. FJ869900 to FJ869905) using the primers described by Hilpertshauser et al. (2006). The most divergent isolate of the Mozambican isolates, Chobela1, was amplified from two different samples, and shows three mismatches and 99% identity with several previously submitted sequences in a BLAST search. The isolates Chobela2, Chobela3 and Chobela4 are 99% identical with a *B. bigemina* isolate identified in China (GenBank accession no. AY603402). Chobela5 and

Chobela6 are both 99% identical with isolates from Kenya (GenBank accession no. EF458200) and Argentina (GenBank accession no. EF458191). An alignment with the *B. bigemina* 18S ribosomal RNA gene sequences and the *B. bigemina* RLB probe was constructed (Fig. 3.2). The sequences from the Mozambican isolates have three mismatches in the region of the species specific probe used in the RLB assay.

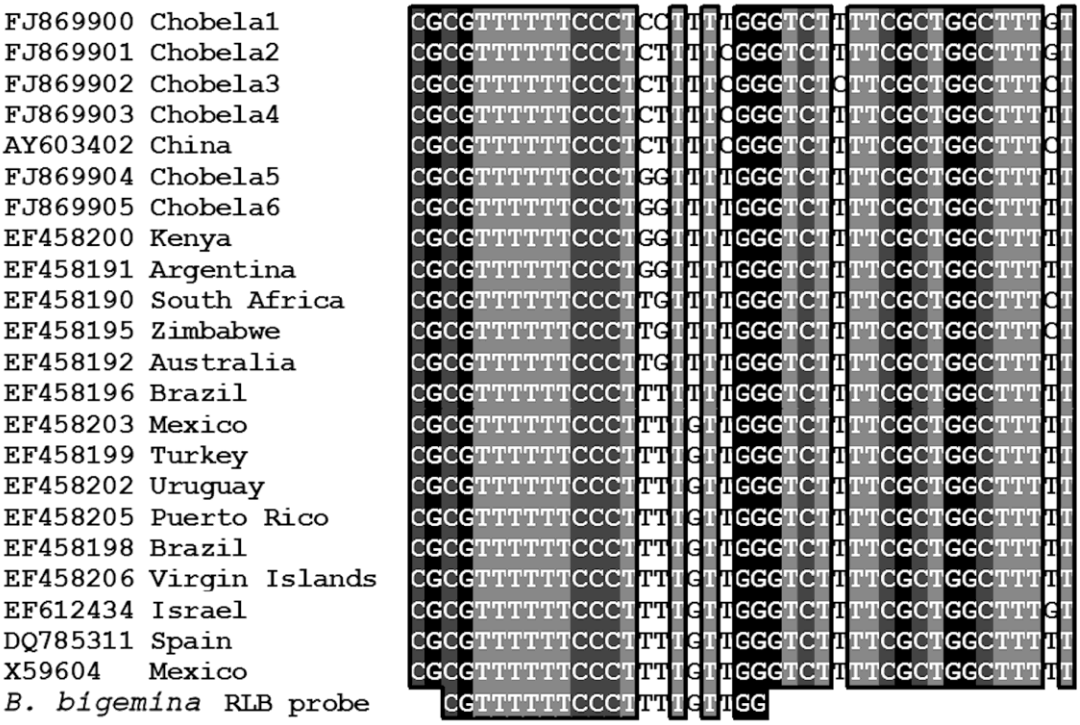

Figure 3.2. Alignment of worldwide *B. bigemina* 18S rRNA gene sequences with the respective species-specific RLB probe. GenBank accession nos. provided next to the locations of isolates.

Several sequences of the 18S rRNA gene from *Theileria* spp*.* were also amplified and cloned using the primers described by Birkenheuer et al. (2003). Amplified from two samples, *Theileria* sp. MSD/Chobela (GenBank accession no. FJ869895) is 99% identical with *Theileria* sp. "strain MSD" from South Africa (GenBank accession no. AF078816). Two similar isolates, *T. velifera* Chobela1 and Chobela2 (GenBank accession nos. FJ869896 and FJ869897), show 99% identity with *T. velifera* from Tanzania (GenBank accession no. AF097993). Two other homologous isolates, *T. mutans* Chobela1 and Chobela2 (GenBank accession nos. FJ869898 and

FJ869899), are 99% identical with *T. mutans* from Kenya (GenBank accession no. AF078815). A phylogenetic tree with the *Theileria* spp. sequences was constructed (Fig. 3.3).

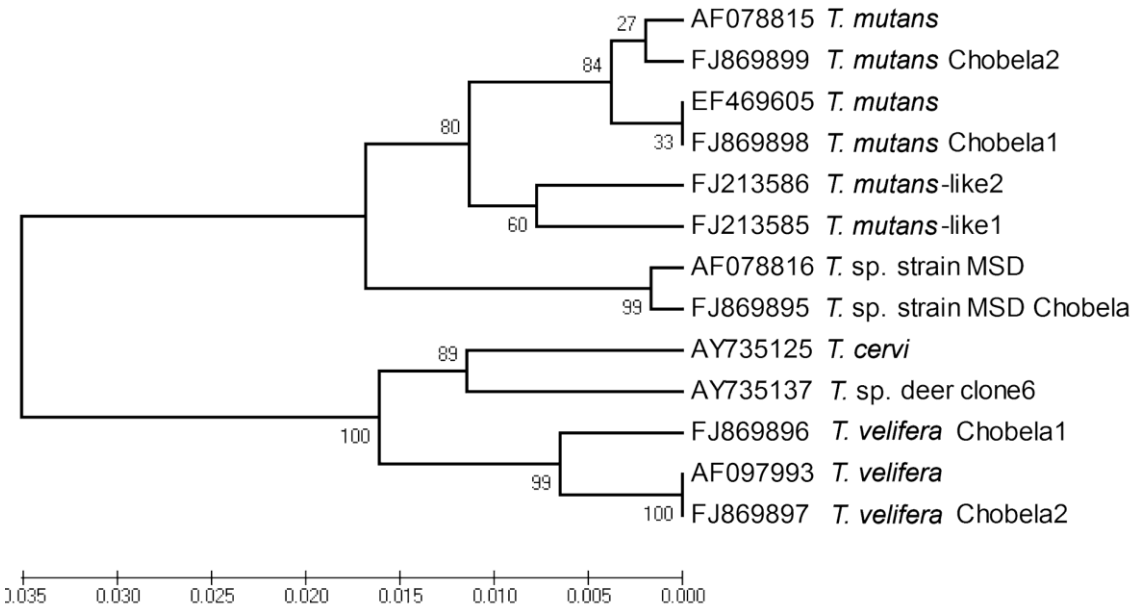

Figure 3.3. Phylogenetic tree inferred from *Theileria* spp. 18S rRNA gene sequences. Tree constructed using the UPGMA method (Sneath and Sokal, 1973), with superimposed bootstrap values (Felsenstein, 1985). GenBank accession nos. provided next to the names.

3.4. Discussion

In this study, the prevalence and the distribution of bovine babesiosis infections are reported for cattle from Maputo province (26,058 km^2) which shares borders with the countries Swaziland and South Africa, being the latest the main origin of imported cattle. Both bovine babesiosis infecting parasites, *B. bigemina* and *B. bovis*, were detected by the seminested hot-start PCR in high frequency in four of the five locations studied. Detection of infection with *B. bigemina* ranged between 30% and 89% (61% total), and infection with *B. bovis* between 27% and 82% (53% total). The results obtained in Maputo province are comparable to the serological results obtained in Tete province where 39% of all calves were seropositive to *B. bovis*, raging between farms from 17% to 71% (Alfredo et al., 2005). Since the five farms are geographically well distributed in Maputo province,

we can conclude that these *Babesia* spp. are common and endemic in this region of southeast Africa although there are low prevalence areas. It was estimated that when the majority of calves are exposed to *B. bovis* infection (at least 75%) by 6 to 9 months of age the disease incidence is very low and a state of natural endemic stability is present (Mahoney, 1974). In all the farms except for the Umbelúzi farm, all the detection values by the seminested hot-start PCR for *Babesia* spp. were below 75%. Therefore the results suggest that bovine babesiosis in Mozambique is not in a situation of endemic stability.

A higher prevalence of infections with *Babesia* spp. was observed in the Umbelúzi farm, and this may be related to the breed of cattle. The European Friesian breed at Umbelúzi is considered more susceptible than the African Nguni breed and Nguni crosses (Bock et al., 2004) present in the remaining locations. Detection frequency of cattle with *Babesia* spp. was above 50% in Maputo Province, whereas in the farm of Bela Vista, both species were detected with the lowest frequency of around 30%. In contrast with the other farms, Bela Vista has a higher relative humidity since it is located in the coastland, in the delta of Maputo River. A higher relative humidity contributes positively to the activity and longevity of the ticks as it prevents desiccation (Sutherst et al., 2006), and this appears contrary to what was observed. Other factors may be contributing to a lower frequency of babesiosis in Bela Vista. In this farm there are also current trypanosomiasis cases, and many drugs used against *Trypanosoma* spp. also show activity against babesiosis (Mamman et al., 1993; Zintl et al., 2003). In Bela Vista, samples were collected in late July, and the collection in other locations was done more than one month later, and this may also have contributed to differences in prevalence of infection with *Babesia* parasites, since babesiosis infections can change on a seasonal basis (Bock et al., 2009). Another exception to the overall results was observed in Namaacha. Only in this farm, *B. bovis* (68%) was detected in higher frequency than *B. bigemina* (55%). Namaacha is located at moderate altitude (approximately 600 meters above sea level), with temperatures lower than those of the other farms, all located at sea level. This environmental factor may contribute to a major change in the challenge with tick species and in consequence with *Babesia* spp., but further studies are needed to confirm this hypothesis.

In acute babesiosis, *B. bovis* infections have lower parasitaemias when compared with *B. bigemina* (Shkap et al., 2005). In the carrier or persistent state, parasitaemia can decrease below 10^{-8}% for *B. bovis* (Calder et al., 1996), and the level of parasitaemia in *B. bigemina* is probably higher than that of *B. bovis* and *B. divergens* (Gubbels et al., 1999). Our results, showed that the seminested hot-start PCR improved the sensitivity over the first hot-start PCR. And it improved more the detection of *B. bovis* (11%) than of *B. bigemina* (2%). This result appears to confirm previous indications that *B. bovis* is present in lower parasitaemias, since only the more sensitive seminested hot-start PCR allowed the detection of infection by *B. bovis* in a significant number of samples.

The haemoparasites detected in higher frequency by the RLB were *T. mutans* (52%) and *T. velifera* (54%). Similar results were obtained by Oura et al. (2004) in a survey conducted in Uganda. In that study, bovine babesiosis was detected by the RLB in low numbers, with 2% for *B. bigemina* and 0% for *B. bovis*. In the present study, *B. bovis* infections were detected by the RLB between 0% and 17%, and *B. bigemina* infections were not detected by the RLB technique, in contrast with the high frequency of detection by the seminested hot-start PCR. Even if, Gubbels et al. (1999) demonstrated that there was no PCR competition in the RLB assay between genomic DNA of *B. bovis* and *T. annulata*, field samples are much more complex with several parasites present at the same time, and with varying proportions between them. In this study, it was found that five different species of the genera *Theileria* and *Babesia* infected the same host simultaneously. Therefore, it is possible that PCR competition occurs and this may explain in part the low sensitivity of the RLB assay compared do the seminested hot-start PCR. To better understand the divergent results obtained by the two methods we decided to look at the local *B. bigemina* 18S rRNA sequences. These results suggest another explanation for the low sensitivity of the RLB. In fact, the analysis of the sequences from the 18S rRNA gene from *B. bigemina* revealed that none of the Mozambican isolates were identical to the oligonucleotide probe described by Gubbels et al. (1999) and used to identify *B. bigemina* in the RLB assay (Fig. 3.2). All the 18S rRNA gene sequences obtained in this study show three mismatches with the oligonucleotide RLB probe. This degree of dissimilarity is enough to prevent the

hybridization of the amplified products from *B. bigemina* with the RLB probe and to originate false negatives, as three mismatches also allows differentiation between species in the RLB assay (Molano et al., 2004; Nindl et al., 2007). Isolates from China (GenBank accession no. AY603402), Kenya (GenBank accession no. EF458200) and Argentina (GenBank accession no. EF458191) are identical to the Mozambican isolates in the probe region (Fig. 3.2). Recently, Petrigh et al. (2008) reported the development of a new probe that allowed the detection of all Argentinean isolates studied, in contrast with the old probe described by Gubbels et al. (1999) that failed in the detection of half of the samples.

In Chobela, 3% of the cattle were positive in the RLB assay for infections with *B. vogeli*, which is known to infect canids, and is transmitted by *Rhipicephalus* (Uilenberg, 2006). To our knowledge, this is the first time that *B. vogeli* was identified in cattle. Tentative attempts to amplify the *B. vogeli* 18S rRNA gene were unsuccessful, although a primer with the sequence of the RLB probe was used besides the ones described earlier. The products that were cloned and sequenced in these tentative amplifications were reported in this study or were identified as similar to *Bos tauros*. Without the confirmation from the amplification of the 18S rRNA gene, the question remains if the positive samples for *B. vogeli* by the RLB assay are in fact from infections by this canids *Babesia* or from a similar unidentified species, or even if it is the result of unspecific hybridization.

The phylogenetic tree inferred from the *Theileria* spp. 18S rRNA genes (Fig. 3.3) shows the taxonomic separation between the *T. mutans* and *T. velifera* groups. None of the *T. mutans* isolates was identical to the isolates from Kenya (Chae et al., 1999) and Sudan (Salih et al., 2007). This may be explained by microheterogeneity where minor nucleotide sequence differences occur in homologous regions of the gene (Chae et al., 1999). In this study, a Chobela isolate (GenBank accession no. FJ869895) was identified with 99% identity with *Theileria* sp. "strain MSD" (GenBank accession no. AF078816), that was isolated from cow blood in South Africa (Chae et al., 1999). These isolates are distinct from *T. mutans*, and *T. mutans*-like1 and 2 isolated from the African buffalo (*Syncerus caffer*). Further studies are needed to clarify the identity and extension of the geographic distribution of this distinct *Theileria* sp. "strain MSD". *T. velifera* is clustered together with the benign

parasite of deer *T. cervi*, and another *Theileria* sp. isolated from deer (Fig. 3.3). The isolates of *T. velifera* identified in this study are not identical to the previous reported sequence from Tanzania (Gubbels et al., 1999), reflecting the microheterogeneity characteristic of *Theileria* spp., and also observed in this study for *T. mutans*.

The results presented in this study, show that multiple infections with several species of cattle haemoparasites are common in Maputo province and that bovine babesiosis is common and endemic in this region, and also suggest high variability within the species in the field. The epidemiological data reported here is indispensable for the assessment of the situation and to establish adequate control measures in Mozambique.

Chapter 4

Identification of papain-like cysteine proteases from *Babesia bigemina* as potential drug targets

4.1. Introduction

The heamoprotozoan parasites of the genera *Babesia* and *Theileria* are tick-transmitted piroplasms in the phylum Apicomplexa (Bock et al., 2004; Chae et al., 1999; Robinson, 1982; Uilenberg, 2006). *Theileria* species have a pre-erythrocytic development stage and invade lymphocytes first in their vertebrate hosts (Spooner et al., 1989). This schizogony is absent in *Babesia* species that only invade erythrocytes (Homer et al., 2000). The cattle pathogens *Babesia bovis* and *B. bigemina* responsible for bovine babesiosis, and *Theileria annulata* and *T. parva* responsible for bovine theileriosis are of veterinary importance as these tick-borne diseases cause severe economic losses throughout the world (Bock et al., 2004; Mukhebi et al., 1992). Live attenuated vaccines are used for the control of cattle babesiosis and theileriosis in many regions of the world, but outbreaks are common as the vaccines are strain-specific (Pipano and Shkap, 2000; de Waal and Combrink, 2006; Yin et al., 2008), suggesting major diversity within the species.

Cysteine proteases are essential to the life cycle and pathogenicity of parasitic protozoa (Sajid and McKerrow, 2002) and are under study as promising drug targets. Cysteine proteases have numerous roles in Apicomplexa parasites such as catabolic functions, immunoevasion, enzyme activation, cellular invasion and rupture (Rosenthal, 2004; Dowse and Soldati, 2004). In piroplasms the role of cysteine proteases is mostly unknown, but the importance of these proteolytic enzymes in the life cycle of *Babesia* and *Theileria* parasites was demonstrated in inhibition studies. It was showed that the lipophilic cysteine protease class inhibitors E-64d and ALLN reduced *in vitro* the invasion of erythrocytes as well as the growth of *B. bovis* (Okubo et al., 2007). The growth of *T. parva* schizont-infected bovine lymphocytes was also impaired by addition of Z-Leu-Val-Gly-CHN_2, a diazomethel ketone inhibitor of CPs (Nene et al., 1990). And, *Theileria (Babesia) equi* propagation of infected erythrocytes *in vitro* was significantly reduced in the presence of the cysteine protease inhibitor E-64d (Holman et al., 2002). Hence, cysteine proteases of piroplasms retain their importance as potential drug targets for the development of reliable and safer chemotherapeutic treatments that are currently required for the control of babesiosis and theileriosis (Vial and Gorenflot, 2006; Wilkie et al., 1998).

Complete piroplasms genome sequences for *B. bovis*, *T. annulata*, and *T. parva* have been reported (Brayton et al., 2007; Gardner et al., 2005; Pain et al., 2005), and sequencing of the *B. bigemina* genome is in active finishing (http://www.sanger.ac.uk/Projects/B_bigemina/). Comparative analysis of these complete genomes revealed a high conservation of gene sequences and synteny, although rearrangements have occurred, and also showed that there are unequally expanded gene families and species-specific genes (Brayton et al., 2007; Pain et al., 2005). Since cysteine proteases roles in protozoa parasites are often related with cell invasion and rupture (Dowse and Soldati, 2004; Rosenthal, 2004), differences in cysteine protease genes content between *Babesia* and *Theileria* genomes may partially explain the ability of *Theileria* species to infect lymphocytes and that *Babesia* species only infect erythrocytes in the vertebrate host. Similar cysteine protease genes may have similar functions in both genera, and some cysteine protease genes may have several functions, as in other Apicomplexa parasites, like with falcipain-1, a papain-like cysteine protease of *P. falciparum* that participates in haemoglobin degradation and in the merozoite release (Dasaradhi et al., 2005).
Here we report the identification of cysteine protease genes of Family C1 from *B. bigemina* and compare them with the annotated genes from the complete piroplasms genomes of *B. bovis*, *T. annulata*, and *T. parva*. We also expressed babesipain-1, one of the identified cysteine proteases from *B. bigemina*, and show that is a proteolytically active enzyme.

4.2. Materials and methods

4.2.1. Database search for similarity and alignment approaches

A total of 2597 query protein sequences of all the annotated members of the C1 family of papain-like cysteine proteases, retrieved from the non-redundant library of protein sequences (pepunit.lib) available at MEROPS (Rawlings et al., 2008) were used in local tblastn search for cysteine protease coding genes in the *B. bigemina* genome contigs database (dated 2008-06-26) available at The Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/Projects/B_bigemina/). A threshold of E≤1e-04 was adopted for the tblastn searches (Wu et al., 2003). The

gene structure of identified cysteine protease genes with multiple exons was predicted using the FGENESH and FGENESH+ software (http://www.softberry.com). Global alignment of homologous chromosomes and contigs used LAGAN (Brudno et al., 2003) with default parameters anchored by predicted gene structures producing a VISTA plot (Frazer et al., 2004). Multiple protein sequence alignment between Piroplasmida CPs was performed using ClustalW (Thompson et al., 1994), followed by manual editing in Bioedit (Hall, 1999). Phylogenetic relationships were constructed and visualized by MEGA4 program (Tamura et al., 2007). The similarities and identities reported for an all-against-all comparison between homologs were determined by MatGAT v2.0 software using the default settings (Campanella et al., 2003).

4.2.2. Cloning and expression of babesipain-1

The full length babesipain-1 gene was amplified from *B. bigemina* genomic DNA (kindly provided by Dr. Varda Shkap from Kimron Veterinary Institute, Israel) by PCR using *Taq* DNA polymerase (Bioron GmbH, Ludwigshafen Germany), forward (BigCis1EX1- G*CCCGGG*AATGTCGGGAACCCGTTCGTACATG) and reverse (BigCis1EX2- G*CTCGAG*TTAGTGGGCCAGCACTGCGCCGGTT) primers which contained *Sma*I and *Xho*I sites (in italics) at their respective 5P-ends. The amplification was performed for 30 cycles of: 94 ºC for 1 min; 55 ºC for 1 min; 72 ºC for 1 min; prior to a final elongation step of 72 ºC for 10 min. The resultant 1392 bp product was ligated into the pTZ57R cloning vector (Fermentas, EU) and several clones were sequenced by STAB Vida (Oeiras, Portugal). Babesipain-1 nucleotide sequence was deposited at GenBank under the accession no. FJ859910.

Full-length babesipain-1 was released from pTZ57R-BigCis1 by digestion with *Sma*I and *Xho*I and subcloned into pGEX-6P-1 expression vector (GE Healthcare, Buckinghamshire UK) digested with the same enzymes. Production and purification of recombinant glutathione S-transferase (GST) fusion protein was conducted according to the manual recommendations (GE Healthcare, Buckinghamshire UK). Briefly, the pGEX-6P-BigCis1 construct was transformed using the TransformAid^{TM} Bacterial Transformation System (Fermentas, EU) into *E. coli* BL21 (GE Healthcare, Buckinghamshire UK), Rosetta-gami^{TM} 2 (Novagen, Madison USA) and XL1-Blue (Stratagene, USA) cells, and cultures were induced at an Abs 600nm of 1.0 for 2h with

1mM isopropyl-β-D-thiogalactopyranoside (IPTG). The induced cell pellet was collected and lysis was performed by sonication. The soluble recombinant protein was purified from the lysate using Glutathione Sepharose 4B (GE Healthcare, Buckinghamshire UK), following the brand indications with the exception that 0.1% Triton X-100 was used in binding and washing steps. Fractions containing enzyme activity were pooled and concentrated by membrane filtration (10 kDa cut-off Vivaspin; Sigma-Aldrich).

Total extracts and purification fractions were analyzed by western blotting for the presence of GST and GST-fusion proteins using Anti-Glutathione-S-Transferase (GST)–Alkaline Phosphatase Conjugate antibody (Sigma-Aldrich) as recommended by the manufacturer.

4.2.3. Activity assays

Babesipain-1 activity was assayed fluorimetrically as the hydrolysis of the substrates Z-Phe-Arg-AMC (Sigma-Aldrich, St. Louis USA), Z-Leu-Arg-AMC, Z-Leu-Leu-Arg-AMC and Z-Val-Val-Arg-AMC (Bachem Holding AG, Bubendorf Switzerland) was followed over time. 20 µl of enzyme solution was added to 100 mM sodium acetate (pH 5.7) containing 50 µM fluorogenic substrate and 5 mM DTT in a final volume of 0.2 ml, and the release of 7-amino-4-methyl coumarin (AMC) was monitored (excitation 360 nm, emission 460 nm) for 30 min at 37 ºC using a Fluorescence Microplate Reader FLx800 (BioTek Instruments Inc., USA).

For gelatin zymography, samples were mixed with SDS-PAGE sample buffer lacking a reducing agent, incubated 5 min at 25 ºC and separated in a 10% (w/v) SDS-PAGE polyacrylamide gel, co-polymerized with 0.1% (w/v) gelatin (Heussen and Dowdle, 1980). The gel was washed twice with 2.5% (v/v) Triton X-100 to remove SDS for 30 min at 25 °C (Liotta and Stetler-Stevenson, 1990). Gel was then incubated overnight at 37 °C in 100 mM sodium acetate (pH 5.7), 5 mM DTT, before staining with Coomassie Brilliant Blue R250. The activity band was observed as a clear colourless area depleted of gelatin in the gel against the blue background.

4.3. Results

4.3.1. Database search for similarity, genome alignments and protein sequence analysis

Five putative genes, named BbiCP1 to 5, with high homology to the query protein sequences of members of the C1 family of cysteine proteases, were found by tblastn in the unfinished *B. bigemina* genome database. Three are intronless genes (BbiCP1 to 3), and the remaining two genes have several exons (BbiCP4 and 5). BbiCP1, 2 and 3 have expected ORF sizes of 1377 bp (identified in Contig_0000353), 1272 bp (Contig_0000413) and 1593 bp (Contig_0000409) respectively. The gene structures of the multiple exon genes were predicted using FGENESH. BbiCP4 was predicted to have nine exons and an ORF size of 1743 bp (Contig_0000381). The ortholog of BbiCP4 in *B. bovis* genome (GenBank accession no. XM_001608666) only was predicted with seven exons and an ORF size of 1641 bp. BbiCP4 has one more upstream exon, and the last two exons resulted from the division of an ancient exon, comparing with the predicted *B. bovis* ortholog. The predicted gene structure of four exons of BbiCP5 and an ORF size of 1563 bp (Contig_0000030) is similar to the four exons and 1626 bp of the annotated *B. bovis* ortholog (GenBank accession no. XM_001609496).

Two annotated proteins in the *T. parva* genome (GenBank accession nos. XP_764234 and XP_764233), showed high identity respectively with the N- and C-terminal of *T. annulata* SERA (TaSERA), but none of them was predicted with the active site cysteine residue. We used FGENESH+ with the amino acidic sequence of TaSERA as template to predict the structure of the SERA gene from *T. parva* (TpSERA) that contained the active cysteine residue, and used this sequence in the subsequent analyzes.

In order to investigate the molecular evolution events of the families of cysteine protease genes from four species of piroplasma parasites, *B. bigemina*, *B. bovis*, *T. annulata* and *T. parva*, genome alignments were conducted using LAGAN and VISTA programs. We can observe that the adjacent regions to the cysteine protease genes are conserved (Fig. 4.1). *T. annulata* genome has seven CP1 paralog genes in tandem and disposed head to tail. In *T. parva* there are only six CP1 genes, and as the result of the difference between these two species of bovine theileriosis a gap in the genome alignment was observed (Fig. 4.1A).

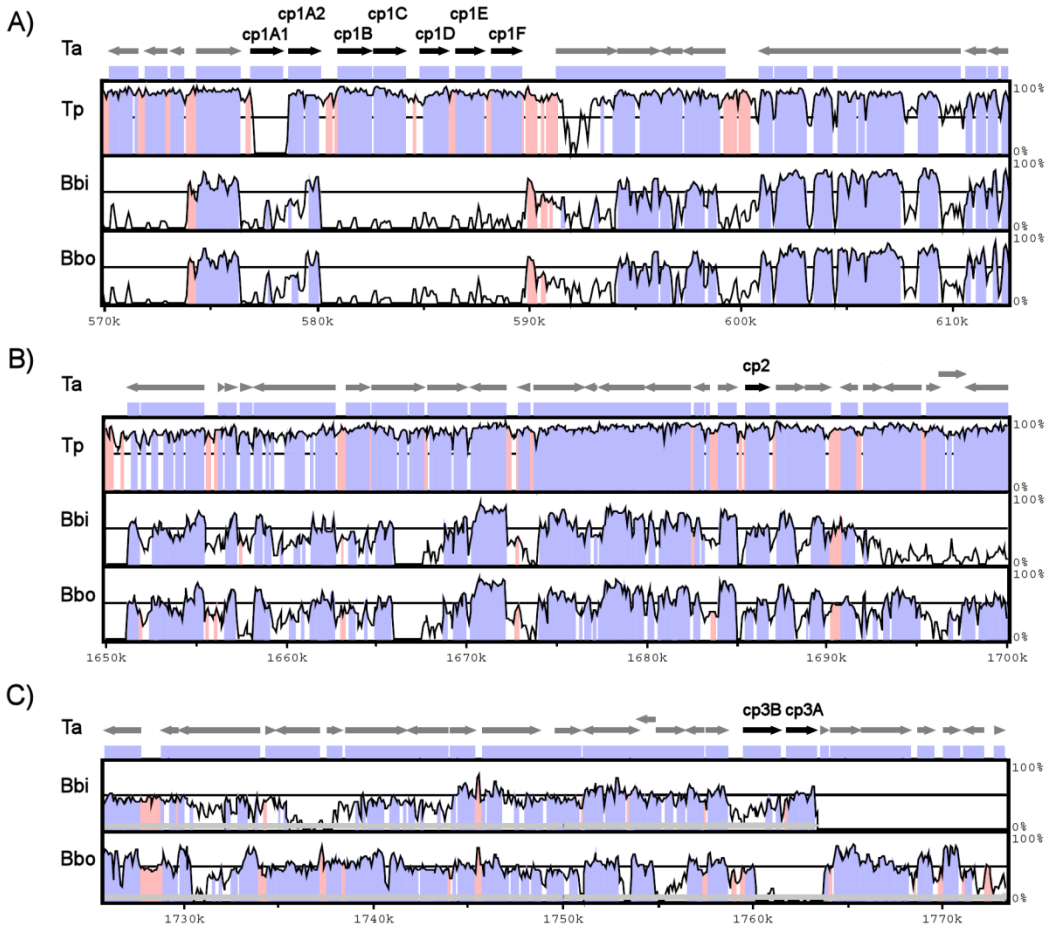

Figure 4.1. VISTA identity plots between homologous regions of Piroplasmida species genomic DNA displaying cysteine protease genes. The gray and black arrows above each VISTA correspond to *T. annulata* annotated exon positions. The level of conservation (vertical axis) is displayed in the coordinates of the *T. annulata* sequence (horizontal axis). Conserved regions above the level of 80%/100 bp (between *Theileria* spp.) or 40%/100 bp are highlighted under the curve (untranslated regions in red and codifying regions in blue). **A)** Identity plot showing cysteine protease 1 (CP1) conserved nucleotide sequences between *T. annulata* (Ta) chromosome 3 (NC_011100), *T. parva* (Tp) chromosome 3 (NW_876245), *B. bigemina* (Bbi) Contig_0000353 and *B. bovis* (Bbo) chromosome 4 (NW_001820857). **B)** Identity plot showing cysteine protease 2 (CP2) conserved nucleotide sequences between *T. annulata* chromosome 2 (NC_011099), *T. parva* chromosome 2 (NC_007345), *B. bigemina* Contig_0000413, and *B. bovis* chromosome 3 (NC_010575). **C)** Identity plot showing cysteine protease 3 (CP3) conserved nucleotide sequences between *T. annulata* chromosome 3 and *B. bigemina* Contig_0000409, and the absence of CP3 in *B. bovis* chromosome 3 (*T. parva* alignment not shown for clearness).

In the *Babesia* species studied there is only one CP1 gene with the same orientation as the orthologs from *Theileria*. There was no duplication or rearrangement events after speciation from a common ancestor in the case of the CP2, CP4 and CP5 gene families (Fig. 4.1B; results not shown for CP4 and CP5). In the CP3 gene family there were duplications in the *Theileria* species. *T. annulata* has two genes (TaCP3A and 3B) in tandem and head to tail (Fig. 4.1C), and *T. parva* has three genes disposed in the same manner (TaCP3A, TaCP3B1 and TaCP3B2). *B. bigemina* genome only has one CP3 gene, BbiCP3, aligned with TaCP3A, although we can observe in *B. bigemina* an untranslated region with moderate homology to the TaCP3B gene (Fig. 4.1C).

In *B. bovis* there is also an untranslated region with homology to TaCP3B but there is no homology to TaCP3A or BbiCP3, in result from the complete deletion of the CP3 like gene. The deletion is in fact a substitution as the length of this region did not changed from *T. annulata* to *B. bovis*. But it seems that it did not resulted in a gene substitution as this putative untranslated region of *B. bovis* in a blastx search against the non redundant protein database at NCBI blast server gave no significant hits. With the deletion of the CP3 like gene, *B. bovis* has four Family C1 cysteine protease genes instead of the five present in *B. bigemina*.

The amino acid sequences of BbiCP1 to 5 were aligned with all known non redundant cysteine protease sequences from other Piroplasmids, and a phylogenetic tree was constructed. We can observe in the tree (Fig. 4.2) that all known CP sequences from Piroplasmids are separated in six main clusters, CP1 to CP5 and SERA. The main differences between *Theileria* spp. and *Babesia* spp. in cysteine proteases occurrence are the absence of the SERA representatives in *Babesia* spp. and duplications in CP1 and CP3 gene families in *Theileria* spp.. Although CP3 is absent from the *B. bovis* genome, it is present in *B. bigemina*, *B. divergens* and *B. caballi*. From the analysis of *T. annulata* and *T. parva* genomes (Fig. 4.1) we observed that gene duplications occurred in the CP1 and CP3 gene families, and from what we can observe in the phylogenetic tree, these duplications also occurred in CP1 gene family in *T. equi* and *T. orientalis*, and in CP3 gene family in *T. cervi* (Fig. 4.2).

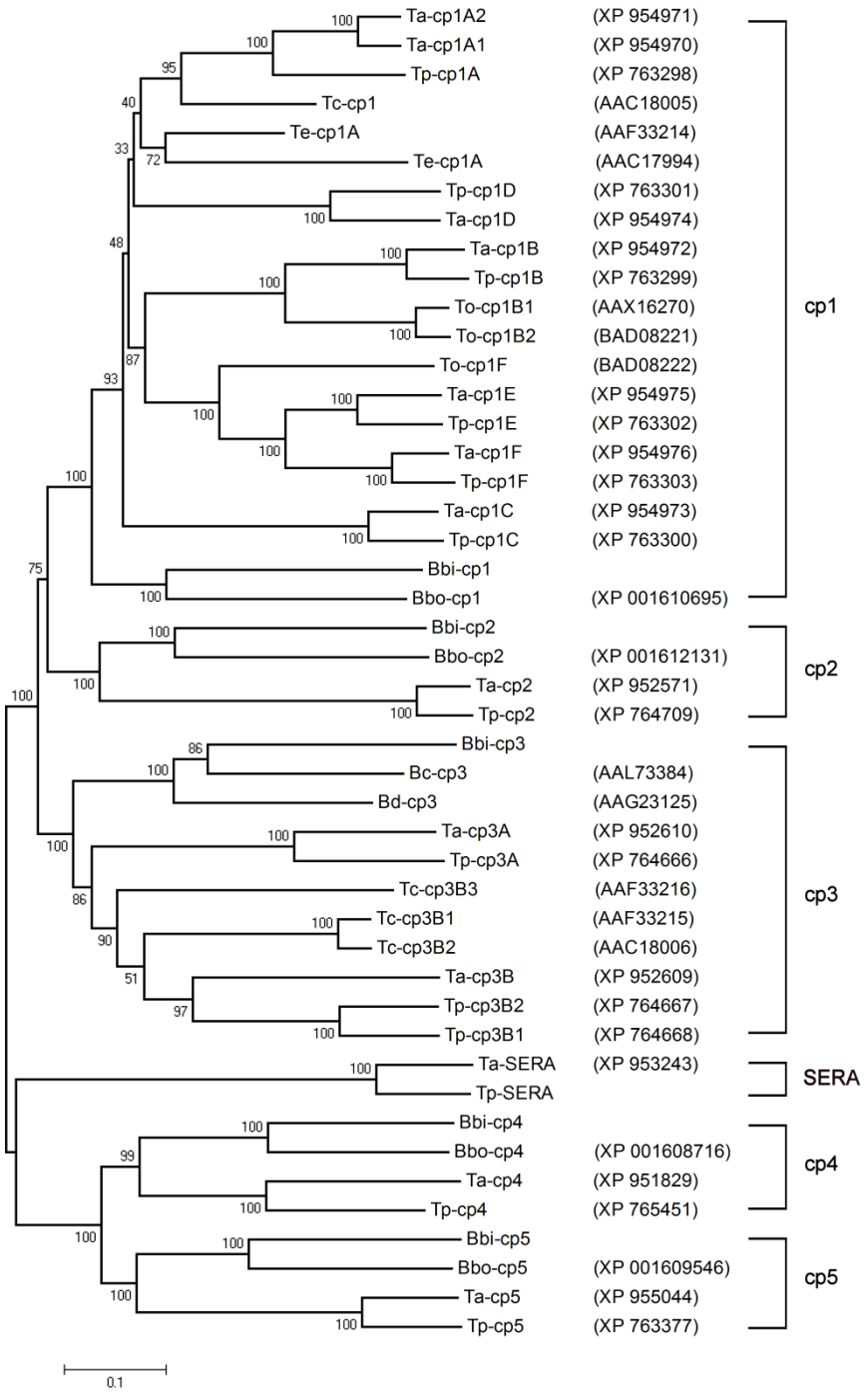

Figure 4.2. Phylogenetic tree inferred from piroplasma cysteine proteases (cp) amino acid sequence data. Piroplasma C1 family cysteine proteases are divided in six main clusters that preceded speciation. Tree constructed using the Neighbor-Joining method (Saitou and Nei, 1987), with superimposed bootstrap values (Felsenstein, 1985). Accession numbers are provided when available. Abbreviations: Bbi, *B. bigemina*; Bbo, *B. bovis*; Bc, *B. caballi*; Bd, *B. divergens*; Ta, *T. annulata*; Tc, *T. cervi*; Te, *T. equi*; To, *T. orientalis*; Tp, *T. parva*.

CP1 genes from *Theileria* spp. have in general a higher identity to their orthologs in other species than to their paralogs in the same genome. For example, ToCP1F share 62% nucleotide sequence identity with TaCP1F and TpCP1F and only 52% with ToCP1B1 and ToCP1B2. In the cases that the paralogs present in the same genome share a higher identify than with the orthologs, they were named with the same letter, as for TaCP1A1 and TaCPA2 that share 92% nucleotide sequence identity, and this suggests a recent duplication event.

CP1 members are in general more identical and similar to papain than the other Family C1 cysteine proteases from Piroplasms (Table 4.1). CP1 to 3 and SERA groups show higher protein sequence similarity with cathepsin L-like CPs, and CP4 and 5 groups with cathepsin C-like.

Table 4.1. Protein sequence similarity and identity matrix. The percentage of similarity (in blue) and identity (in red) of the cysteine protease sequences was calculated by the MatGat v2.0 software (Campanella et al., 2003).

%	BbiCP1	BbiCP2	BbiCP3	BbiCP4	BbiCP5	TaCP1A1	TaCP1A2	TaCP1B	TaCP1C	TaCP1D	TaCP1E	TaCP1F	TaCP2	TaCP3A	TaCP3B	TaCP4	TaCP5	TaSERA	Papain
BbiCP1		27	22	19	20	35	35	26	32	34	31	32	25	23	20	17	18	17	26
BbiCP2	42		21	19	20	30	30	22	26	26	29	28	32	21	19	14	18	14	25
BbiCP3	40	37		18	17	24	24	19	24	21	21	21	18	29	25	16	19	13	20
BbiCP4	34	33	36		32	20	20	18	18	20	20	20	18	19	19	32	31	14	15
BbiCP5	36	34	36	49		18	18	17	19	18	18	19	17	20	17	21	36	18	19
TaCP1A1	52	47	38	37	35		92	37	41	39	45	43	22	25	21	16	18	15	25
TaCP1A2	52	47	39	37	35	95		38	43	40	44	42	24	25	21	16	17	15	25
TaCP1B	46	40	38	34	34	54	56		32	34	35	34	20	22	18	14	18	15	21
TaCP1C	50	44	39	35	34	60	61	50		38	37	36	24	26	21	15	17	15	24
TaCP1D	51	46	37	35	37	60	60	52	58		41	38	26	24	20	15	17	14	26
TaCP1E	50	47	36	36	35	63	63	54	58	62		71	27	24	21	15	18	18	25
TaCP1F	50	48	37	33	38	64	63	53	55	60	81		27	26	20	14	19	16	26
TaCP2	42	51	34	31	31	44	44	39	41	48	47	46		23	18	14	15	15	22
TaCP3A	40	37	46	35	34	42	42	41	41	43	42	41	37		30	17	18	16	22
TaCP3B	37	33	45	37	31	35	35	34	36	33	37	36	31	44		19	18	17	19
TaCP4	30	26	29	48	36	29	29	27	30	28	28	29	26	30	31		23	17	15
TaCP5	35	30	36	48	55	36	35	37	35	38	34	35	30	38	34	35		17	18
TaSERA	31	29	27	28	34	30	30	30	32	28	31	30	29	33	33	30	34		15
Papain	40	43	35	31	31	43	42	35	39	45	41	42	37	33	28	23	32	26	

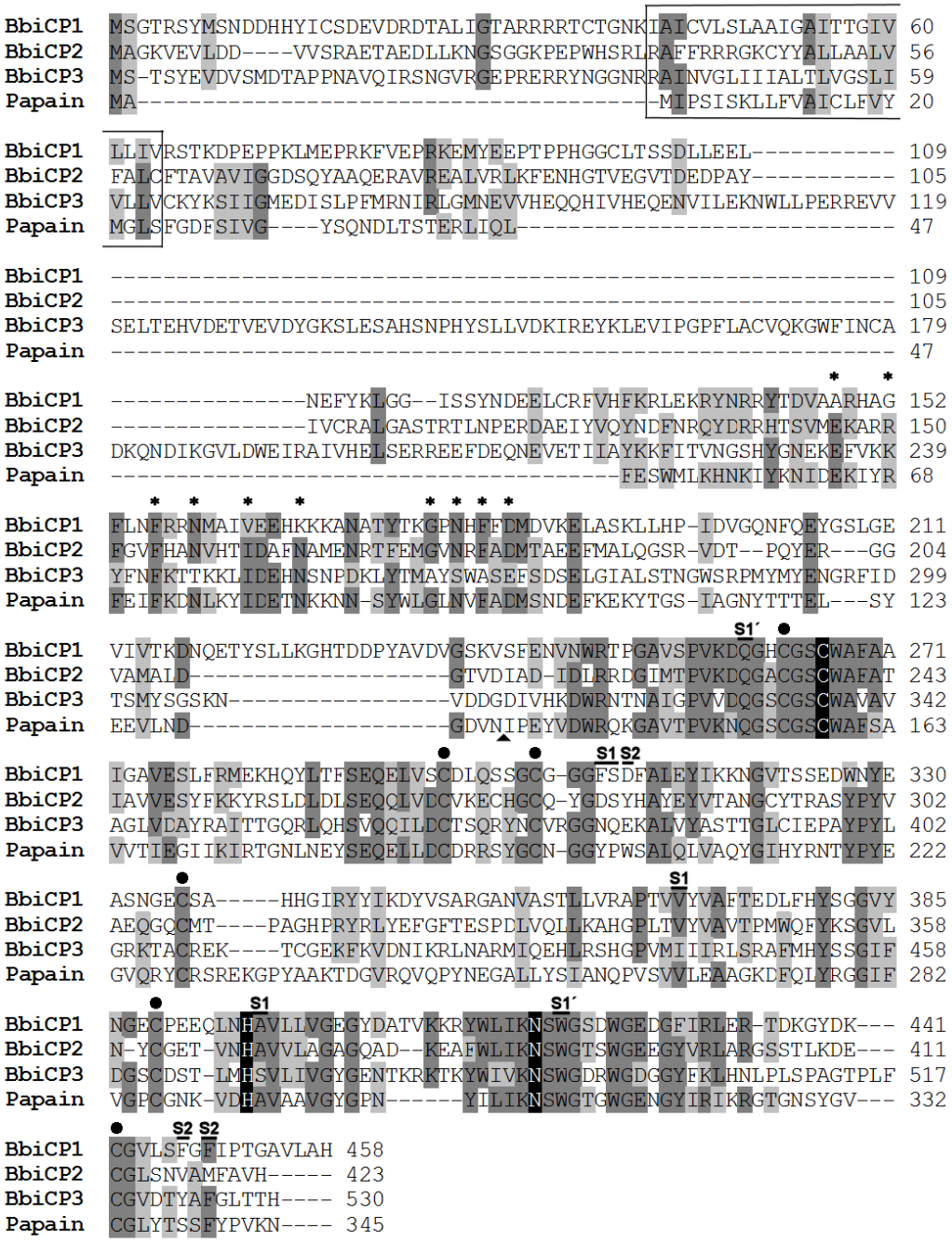

Figure 4.3. Multiple sequence alignment of the deduced amino acid sequences of the cathepsin L-like cysteine proteases BbiCP1 to CP3 and papain. Identities and similarities are shaded (75%) in dark and light grey respectively. Predicted transmembrane domains are boxed, amino acids representing the ERFNIN and GNFD prodomain motifs are labelled with asterisks, the position of the mature domain processing site of papain are indicated by an arrowhead, all cysteine residues involved in disulphide bridges are labelled by a closed circle, residues making up the active sitecleft (S1, and S2 subsites) are indicated accordingly, and catalytic amino acids are shaded in black.

BbiCP1 to 3 shares many features with papain, including a 20-23 amino acid putative trans-membrane domain, the presence of the ERFNIN and GNFD pro-sequence motifs typical of cathepsin L-like cysteine proteases (Sigwali et al., 2001), conservation of catalytic residues, and six cysteine residues involved in disulphide bond formation in the mature protease sequence (Fig. 4.3).

BbiCP1 to 3 have hydrophobic residues (Phe, Val and Tyr respectively) at the amino acid residue 205 (mature papain numbering) of the protease S2 pocket, typical of mammalian cathepsin L-like cysteine proteases (Sajid and McKerrow, 2002). The amino acid residue 205 is thought to be critical to the substrate preference in the S2 subsite, possibly because hydrophobic residues do not stabilize polar residues and thus limiting access to the catalytic triad. BbiCP1 to 3 have an unusual insert between highly conserved amino acid residues near the C-terminus and long prodomains of 200 to 300 amino acids, typical of Apicomplexa cysteine proteases (Rosenthal, 2004).

All the other cysteine proteases from *Babesia* and *Theileria* also have long prodomains. CP1B, CP3A and SERA from *Theileria* spp. have C-terminal extensions with a high percentage of amino acids with electrically charged side chains (Lys, Arg and Glu). In *Babesia* spp. the active site residues are conserved, but in *Theileria* spp. there are some unusual mutations (Fig. 4.4). TaCP2 and TpCP2 have a serine residue instead of the active site cysteine residue, also observed in some members of SERA family from *Plasmodium falciparum* (Hodder et al., 2003). TaSERA and TpSERA do not show this substitution and have the conserved active site cysteine of cysteine proteases. CP1B family have a mutation in the active site cysteine into aspartic acid (TaCP1B and TpCP1B) or into glycine (ToCP1B1 and B2). The majority of the CP1 proteins from *Theileria* spp. have an unusual insertion of one or two residues between the glutamine of the oxyanion hole and the active site cysteine (Fig. 4.4).

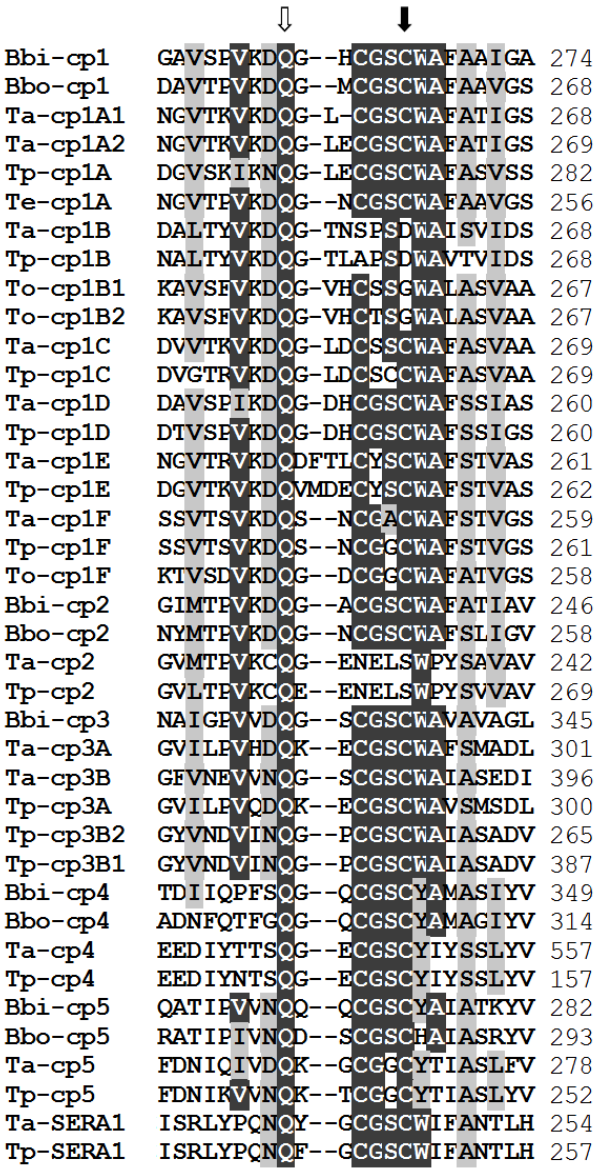

Figure 4.4. Sequence alignment spanning the active site cysteine (black arrow) and the predicted glutamine of the oxyanion hole (white arrow) of cysteine proteases (cp) from *B. bigemina* (Bbi), *B. bovis* (Bbo), *T. annulata* (Ta), *T. parva* (Tp), *T. orientalis* (To) and *T. equi* (Te).

4.3.2. Recombinant expression and activity of babesipain-1

The full length babesipain-1 gene was PCR amplified from genomic DNA. Babesipain-1 nucleotide sequence (GenBank accession no. FJ859910) from an isolate from Israel is 98% identical to the sequence of BbiCP1 that was identified in the *B. bigemina* genome database (virulent Australian isolate). The full length babesipain-1 gene was cloned into the expression vector pGEX-6P-1, and the recombinant fusion protein GST-babesipain-1 was successfully produced in a recombinant *E. coli* system (Fig. 4.5A).

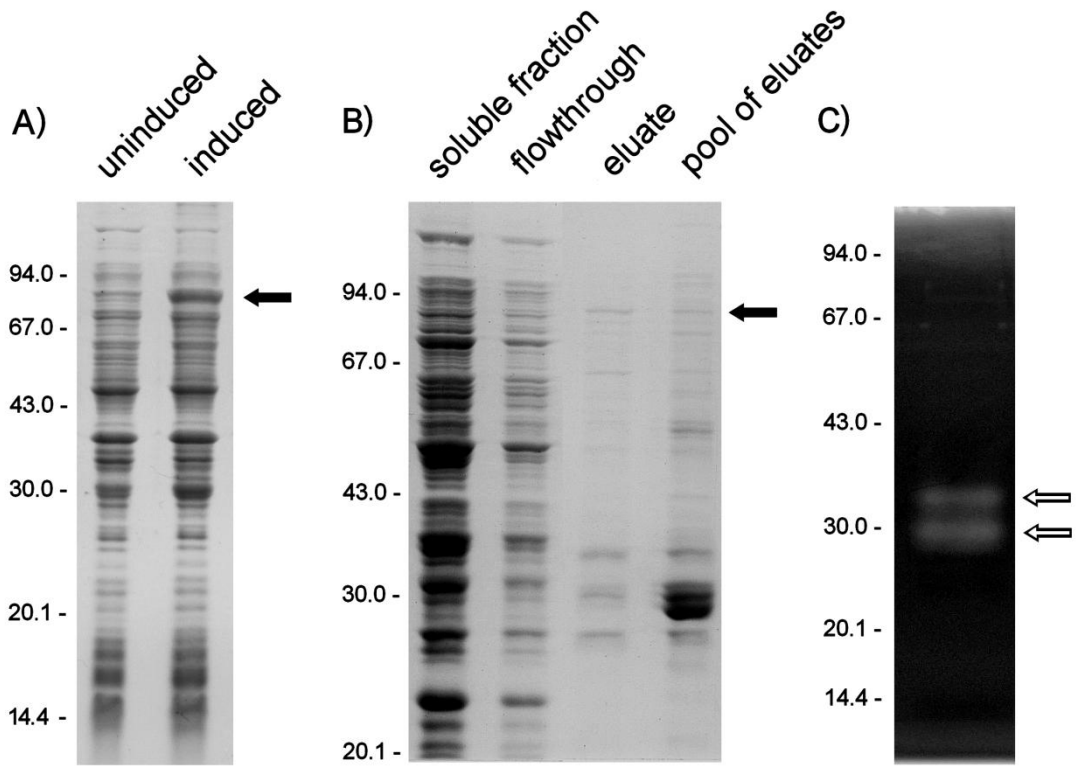

Figure 4.5. Expression, purification and gel substrate activity of recombinant babesipain-1. The expression of the 75 kDa GST-babesipain-1 was evaluated by 12.5 % (w/v) SDS-PAGE (A). Uninduced *E. coli* Rosetta-gami 2 cells and induced with IPTG for 2h. Fusion GST-babesipain-1 protein was purified by affinity chromatography (B): soluble fraction of induced cells, flowthrough, and eluate. The concentrated pool of eluates was also evaluated by gelatin zymography (C) showing activity. The positions of molecular weight markers are labelled in kDa. Black arrows indicate the GST-babesipain-1 fusion protein and white arrows indicate activity.

Three *E. coli* strains were transformed with the pGEX-6P-BigCis1 expression construct, and a small scale expression experiment was conducted to compare expression levels, which were similar in the different strains. The liquid cultures of *E. coli* Rosetta-gami™ 2 (RG2) transformed using TransformAid™ Bacterial Transformation System (Fermentas, EU) with pGEX-6P-BigCis1 had a reduced duplication time (i.e. toxicity) in comparison with cultures of the transformed strains XL1-Blue and BL21 and non transformed RG2. Commercial competent RG2 (Novagen) were also transformed with pGEX-6P-BigCis1, but while colonies were visible in solid LB-Lennox medium (Lennox, 1955), growth was not observed in

liquid cultures. Soluble extracts of the IPTG induced cells pellet of the three strains transformed with the expression vector pGEX-6P-BigCis1 showed protease activity with the four substrates used in the assay (Fig. 4.6A). Soluble fractions of transformed RG2 showed an 11 to 49 fold increase in activity in comparison with transformed BL21 depending on the substrate. Activity was not detected in non transformed cells or transformed with the naked vector pGEX-6P-1. We concluded that the observed protease activity was from the recombinant babesipain-1, and probably it was responsible for the toxicity observed in transformed RG2 cultures.

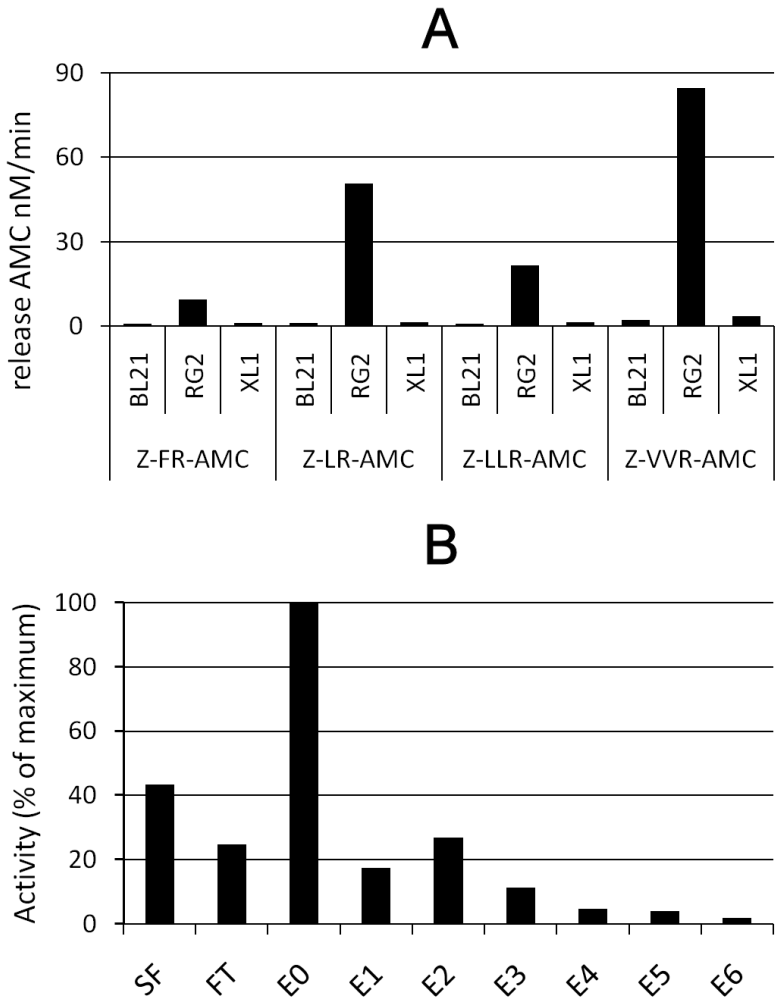

Figure 4.6. Babesipain-1 activity assays. (A) Effect of *E. coli* strain and substrate in the activity of soluble fractions from induced cells with IPTG. (B) Activity of fractions collected in the affinity purification. Abbreviations: RG2, Rosetta-Gami2; XL1, XL1Blue; SF, soluble fraction; FT, flowthrough; E, elution.

The majority of the recombinant GST-babesipain-1 protein was expressed in the insoluble form, since an increase of expression was observed in the total extract after induction (Fig. 4.5A) and the recombinant protein was undetectable in the soluble fraction (Fig. 4.5B). The soluble GST-babesipain-1 fusion protein was purified by affinity chromatography, but although alterations to the protocol were made, the elution fractions showed several proteins in a SDS-PAGE gel (Fig. 4.5B). We assessed the protease activity of the purification fractions (Fig. 4.6B) and observed that the column retained the GST-babesipain-1 fusion protein, since the activity of the pre-elution fraction (Eluate 0) was higher than the activity of the initial soluble fraction. The lower activity of the GSH elution fractions (Eluate 1 to 6) indicates that babesipain-1 retained activity at neutral pH (1x PBS pH 7.3 used in the purification steps) and self processed the GST-propart prior to the elution steps with GSH. Self-processing of GST-babesipain-1 was confirmed by gel substrate activity analysis (Fig. 4.5C) that revealed activity at approximately 30 kDa, the predicted size of the mature enzyme. The presence of two bands in the gel substrate indicates that self-processing was incomplete and is possibly a multi-step process, which explains the presence of several bands in the elutions and in the concentrated pool of the GSH elutions (Fig. 4.5B). The two strong bands observed in the SDS-PAGE gel at 30 kDa and 27 kDa, present in the pool of eluates, corresponds to the released GST protein (confirmed by western blotting), which was resistant to babesipain-1 proteolytic activity. The activity of the partially processed and mature babesipain-1 (pool of the eluates) was completely inhibited by the specific inhibitors of cysteine proteases E-64 (5µM) and leupeptin (50µM).

4.4. Discussion

Here we report the identification in the unfinished *B. bigemina* genome of five putative genes encoding papain-like cysteine proteases of Family C1 and the relationship with genes from the complete and annotated genomes of several piroplasms. We also showed that one of the newly identified genes in the *B. bigemina* genome encodes an active cysteine protease. Babesipain-1 is a cathepsin

L-like cysteine protease that shares many features with papain (Fig. 4.3) and is the first cysteine protease from piroplasms described with proteolytic activity.

Although cysteine proteases usually have autolysis activity at acidic pH (Brömme et al., 2004), babesipain-1 auto processed at neutral pH and also showed activity in the cultures of transformed *E. coli* RG2 cells. Even if *E. coli* RG2 cells promoted the correct folding of a higher amount of babesipain-1 in comparison with the other *E. coli* strains used, the majority of the recombinant protein was expressed as inclusion bodies. The combination of these factors greatly reduced the final yield of recombinant babesipain-1 obtained. Other production and/or purification strategies must now be developed in order to obtain sufficient amounts of recombinant babesipain-1 to proceed to further characterization and inhibition studies. Nevertheless, activity of babesipain-1 was completely inhibited by E-64, and the lipophilic compound but similar in action E-64d effectively inhibit *in vitro* the growth of *B. bovis* parasites and invasion of host erythrocytes (Okubo et al., 2007). In *B. bovis* the probable candidates for E-64d inhibition are the two cathepsins L-like, BboCP1 and BboCP2, and possibly calpain (GenBank accession no. XP_001611895). CP1 (babesipain-1) of *Babesia* spp. is therefore a potential drug target of a shortlist of essential cysteine proteases.

Genes of cysteine proteases of piroplasms were previously described (Baylis et al., 1992; He et al., 2005; Holman et al., 2002; Nene et al., 1990; Sako et al., 1999), as single or multi copy genes but the relationship between them was not clear. In view of the results presented here, previous comparisons were not always made between orthologs, or even between members of the same phylogenetic group, principally due to the lack of information at the time of analysis. We showed that there are six distinct groups of cysteine protease genes of Family C1 in *Theileria* spp. and four to five groups in *Babesia* spp.. More than the absence of the SERA like protease gene group in *Babesia*, the overall content of cysteine protease genes suggests that after the separation from a common ancestor, *Babesia* lost cysteine protease genes and *Theileria* gained through duplication of the existent ones (13 cathepsin L and C-like genes present in both *T. annulata and T. parva* against only 5 in *B. bigemina* and 4 in *B. bovis*). All analysed species have two cathepsin C-like groups, and therefore the differences observed are in the number of cathepsin L-

like groups. Molecular evolution in *Theileria* occurred not only in a sheer increase in the number of genes, but also in sequence diversity. The active sites of cysteine proteases of *Theileria* CP1 and CP2 groups have unusual mutations and insertions that show no parallel in *Babesia* (Fig. 4.4). CP1B of *T. annulata* and *T. parva* show a substitution in the active site cysteine into aspartic acid and, in *T. orientalis,* the apparent proteolytically inactive CP1B has the catalytic cysteine substituted into glycine, and show binding affinity to haemoglobin (He et al., 2005). Although it was suggested that this binding activity could be related to detoxification/uptake of haeme, a mechanism also present in other protozoan parasites like *Leishmania donovani* (Krishnamurthy et al., 2005), this could imply that *Babesia* developed a different mechanism of detoxification and/or feeding since there is only one CP1 gene in *Babesia* with proteolytic activity as we showed. Most of the CP1 group members from *Theileria* spp. have an unusual insertion of one or two residues between the glutamine of the oxyanion hole and the active site cysteine (Fig. 4.4). Of the CP1 members that maintain the catalytic cysteine residue, CP1C and CP1E have unusual mutations of the conserved glycine^{23} (papain numbering) to serine and tyrosine respectively. In the rodent *Plasmodium* spp. a mutation in glycine^{23} to alanine, altered specificity and decreased activity (Singh et al., 2002). It was suggested that this substitution in the rodent *Plasmodium* spp. may confer an increase in specificity to necessary substrates without unwanted proteolytic digestion, and this also may occur in *Theileria* spp.. Also, in another protozoan parasite, *Leishmania mexicana*, the cathepsin L-like *cpb* cysteine protease genes found in allelic tandem arrays of 19 copies, have distinct activities reflecting the differences in the amino acid sequences, that suggested the authors that the isoenzymes codified by the individual genes of the array perform different functions (Mottram et al., 1997). The mutations and insertions observed between the glutamine of the oxyanion hole and the active site cysteine of *Theileria* CP1 and CP2 groups (Fig. 4.4) may therefore confer specificity to different substrates and possibly implicating different roles in the life cycle.

The *Theileria* species studied have a single copy gene similar to the multigene SERA family from *P. falciparum* (PfSERA), and all are adjacent in the respective genomes to a HesB member protein (McCoubrie et al., 2007) that putatively

participates in the biosynthesis of iron-sulphur proteins. SERA family members possess a central domain that has homology to papain. PfSERA5 is a fairly typical cysteine protease except where the active site cysteine has been replaced with a serine (Hodder et al., 2003). A similar mutation also occurred in *Theileria*, but remarkably it occurred in the CP2 gene family. It is interesting that this unusual mutation occurred in different gene families and organisms. The serine-type PfSERA5 possess chymotrypsin-like proteolytic activity, and is thought to be involved in parasite egress and/or erythrocyte invasion (Hodder et al., 2003). The cysteine-type SERAs of *Plasmodium* spp. have primary roles in the pre erythrocytic stage and in the mosquito sexual stages (Aly and Matuschewski, 2005; Schmidt-Christensen et al., 2008). The *Theileria* SERA is likely to be functionally more related to the cysteine-type SERA than to the serine-type SERA from *Plasmodium* spp., since it contains the catalytic cysteine, and therefore may also play a role in the pre erythrocytic stage. It is noteworthy that CP1B, CP3A and SERA from *Theileria* spp. have putative C-terminal extensions that may be cleaved by another protease, similar to the proteolytic maturation what occurs with PfSERA (Yeoh et al., 2007). Further studies are needed to clarify the functional role of *Theileria* SERA and its relationship with PfSERA.

Babesia apparently lost its SERA gene maintaining an ortholog of the HesB member protein (GenBank accession no. XP_001611819) that in *Theileria* spp. and *Plasmodium* spp. is clustered next to the SERA genes (Arisue et al., 2007). In *B. bovis* genome it also occurred a recent deletion of the CP3 gene given that this gene is present in *B. bigemina*, *B. caballi* and *B. divergens* (Fig. 4.2). The biological significance of these deletions remains uncertain, but it is the function of the products of the undeleted genes in *Theileria* that may help us comprehend. *Babesia* do not invade lymphocytes as *Theileria* do, and since cysteine proteases are involved in invasion or egress in other parasitic protozoa (Dowse and Soldati, 2004; Rosenthal, 2004; Sajid and McKerrow, 2002; Yeoh et al., 2007), the total differences observed in CPs between these genera may, in part, explain these rather different mechanisms of infection in the host vertebrates.

In conclusion, although comparisons of the genomes of *Babesia* and *Theileria* species revealed high conservation of gene sequences and synteny (Brayton et al.,

2007; Pain et al., 2005), in respect to cathepsin L-like cysteine proteases, there are significant differences between both genera. *Babesia* spp. have a low number of cathepsin L-like cysteine proteases (2 to 3 genes) in comparison with the closely related *Theileria* spp. (11 genes) and also with *Plasmodium* spp. (16 genes in *P. falciparum*)(Wu et al., 2003; Rosenthal, 2004). The limited and thus essential babesial cathepsin L-like cysteine proteases maintain their importance as potential drug targets.

Chapter 5

Characterization of the potential drug target cysteine protease babesipain-1

5.1. Introduction

Babesiosis has long been recognized as an economically significant disease of cattle (Bock et al., 2004). Live vaccines are the main measure to control bovine babesiosis in many regions of the world, but breakdowns of protective immunity are common as the vaccines do not offer cross protection against the several local existing variants (de Waal and Combrink, 2006; Fish et al., 2008). Chemoprophylaxis is also used as a method of short-term protection, during outbreaks, and other temporary circumstances, as when pregnant cows are at increased risk of transmission and consequent abortion or when animals are momentarily relocated (Bock et al., 2004; de Waal and Combrink, 2006). Extensive outbreaks of the disease also occur in herds of susceptible unvaccinated cattle transported to natural endemic stable areas (Alfredo et al., 2005; Bock et al., 2004). However, control of babesiosis currently requires more specific, fast acting, reliable and safer chemotherapeutic treatments (Vial and Gorenflot, 2006). Thus, the identification and characterization of new drug targets for chemotherapy is a pressing priority.

The members of the Clan CA, or papain-family of cysteine proteases are the most common proteases in protozoan parasites and are essential to the life cycle and pathogenicity of these organisms (Sajid and McKerrow, 2002). Hence, cysteine proteases of protozoan parasites are recognized drug targets and specific inhibitors are in validation for chemotherapy of leishmaniasis, malaria, and trypanosomiasis (McKerrow et al., 2008). In *Babesia bovis*, another agent of bovine babesiosis, essential roles for cysteine proteases were suggested when it was demonstrated that cultured parasites showed reduced growth when incubated with the cell permeable cysteine protease inhibitors ALLN and E-64d (Okubo et al., 2007). Cysteine proteases of *Babesia* spp. may therefore be suitable as new chemotherapeutic targets. Three cathepsin L-like cysteine proteases, named babesipains, were already identified in the *B. bigemina* genome (**Chapter 4**). Babesipain-1 was cloned, expressed as a fusion protein with gluthatione S-transferase (GST) and the soluble protein purified by affinity chromatography. The recombinant babesipain-1 had activity against typical peptide substrates of cysteine proteases, and was inhibited by E-64, but the low yield of the soluble

purification prevented additional characterization (**Chapter 4**). In this chapter, we report the improved production of babesipain-1 by purification of the insoluble recombinant form, refolding of the denatured protein and the biochemical characterization of this potential drug target.

5.2. Materials and methods

5.2.1. Materials

The inhibitors 4-(2-Aminoethyl) benzenesulfonyl fluoride (AEBSF), 3,4-dichloroisocoumarin (DCI), ethylenediaminetetraacetic acid (EDTA), phenylmethylsulfonyl fluoride (PMSF), egg cystatin, 1,10-phenanthroline, N-acetyl-L-leucyl-L-leucyl-L-argininal (leupeptin), pepstatin-A, N-(trans-epoxysuccinyl)-l-leucine 4 guanidinobutylamide (E-64), and N-acetyl-leucineleucine-norleucinal (ALLN) were obtained from Sigma.

5.2.2. Expression and purification of recombinant babesipain-1

The babesipain-1 gene was previously amplified with primers containing appropriate restriction sites and cloned (**Chapter 4**). Full-length babesipain-1 was released from pTZ57R-BigCis1 by digestion with *Sma*I and *Xho*I and subcloned into pGEX-6P-1 expression vector (GE Healthcare, Buckinghamshire UK) digested with the same enzymes. The pGEX-6P-BigCis1 construct was transformed using the TransformAid™ Bacterial Transformation System (Fermentas, EU) into *E. coli* BL21 cells (GE Healthcare, Buckinghamshire UK), and cultures were induced at an Abs 600nm of 1.0 for 3h with 1mM isopropyl-β-D-thiogalactopyranoside (IPTG).

The induced cell pellet from 1 L of culture was collected, resuspended with 30 mL of 100 mM Tris pH 8.0, 0.2 % (v/v) Triton X-100, and lysis was performed by sonication. Lysed cells were diluted to 250 mL in Urea Wash Buffer (100 mM Tris pH 8.0, 0.2 % (v/v) Triton X-100, 500 mM Urea) and stirred for 30 min at room temperature. Insoluble material containing the recombinant protein was pelleted by centrifugation at 27,000 × g for 10 min. The pellet was resuspended with 250 mL Urea Wash Buffer, stirred and pelleted in the same previous conditions. This procedure was repeated 3 more times. The washed pellet was then resuspended in 40 mL solubilization buffer (8 M Urea, 100 mM Tris pH 8.5, 32 mM β-mercaptoethanol) and stirred overnight at 4 ºC

to solubilise the recombinant material. Residual insoluble material was removed by centrifugation at 27,000 × g for 30 min.

5.2.3. Refolding of recombinant babesipain-1

Refolding of the recombinant babesipain-1 was optimized by testing more than 100 different buffer combinations in a microplate format as described previously (Sijwali et al., 2001a). Buffers differed in pH, the concentration of redox couple (reduced glutathione (GSH) to oxidized glutathione (GSSG)), aggregation suppressors (L-arginine, KCl, Triton X-100, and polyethylene glycol), and co-solvents (glycerol and sucrose). Refolding was initiated by 100-fold rapid dilution (Hirose et al., 1989) of ice-cold reduced–denatured protein to a final protein concentration of 0.1 mg/mL in 300 µl of ice-cold refolding buffer followed by incubation at 4 ºC for 16 h. Refolding efficiency was evaluated by assaying 50 µl of each refolding reaction for hydrolysis of Z-Phe-Arg-AMC as described below (with modification in pH to 5.0), and the efficiency for each reaction was defined as the percentage of the maximum difference of hydrolysis achieved in this experiment. For large-scale refolding, recombinant babesipain-1 was rapidly diluted 100-fold in 2 litres of ice-cold optimized refolding buffer (20 mM Tris pH 11, 1 mM EDTA, 5 mM GSH, 0.5 mM GSSG) and incubated at 4 ºC for 16 h. The resulting soluble matter was concentrated 10 times using an Amicon stirred cell (Millipore, Watford, UK) fitted with a 10 kDa cut-off membrane. To allow processing to an active enzyme, the concentrated preparation was acidified by the addition of one-tenth volume of 1 M sodium acetate buffer pH 4.6 (final pH 5.0) and the sample was incubated at 4 ºC for 20 h. Insoluble material was removed by centrifugation at 27,000 × g for 30 min. The resulting soluble protein was then further concentrated 8 times using an Amicon stirred cell fitted with a 10 kDa cut-off membrane and was stored in 50 % (v/v) glycerol at −20 ºC.

5.2.4. Activity assays

Babesipain-1 activity was assayed fluorimetrically as the hydrolysis of the substrates Z-Phe-Arg-AMC (Sigma-Aldrich, St. Louis USA), Z-Leu-Arg-AMC, Z-Leu-Leu-Arg-AMC and Z-Val-Val-Arg-AMC (Bachem Holding AG, Bubendorf Switzerland) was followed over time. 8 µL of enzyme solution was added to 100 mM sodium acetate pH 5.5 containing 50 µM fluorogenic substrate and 1 mM DTT in a final volume of 0.2 mL

(or with changes in pH or reductant as described), and the release of 7-amino-4-methyl coumarin (AMC) was monitored (excitation 360 nm, emission 460 nm) for 30 min to 1 h at 37 ºC with a Fluorescence Microplate Reader FLx800 (BioTek Instruments Inc., USA).

For gelatin zymography, samples were mixed with SDS-PAGE sample buffer lacking a reducing agent, incubated 5 min at 25 °C and separated in a 12.5% (w/v) SDS-PAGE polyacrylamide gel, co-polymerized with 0.1% (w/v) gelatin (Heussen and Dowdle, 1980). The gel was washed twice with 2.5% (v/v) Triton X-100 to remove SDS for 30 min at 25 °C (Liotta and Stetler-Stevenson, 1990). Gel was then incubated overnight at 37 °C in 100 mM sodium acetate (pH 5.5), 1 mM DTT, before staining with Coomassie Brilliant Blue R250. The activity band was observed as a clear colourless area depleted of gelatin in the gel against the blue background.

5.2.5. Enzyme Kinetics

Babesipain-1 enzyme concentration was determined by titration with E-64 (Barrett and Kirschke, 1981). Rates of hydrolysis of fluorogenic peptide substrates were determined in the presence of constant babesipain-1 enzyme concentration (5.13 nM). The kinetic constants K_m and k_{cat} were determined using the GraphPad Prism 5.0 program (GraphPad Software, San Diego California USA). The derived kinetic constants K_m and k_{cat} always had estimated errors of <15%.

5.3. Results

The full length babesipain-1 gene was previously cloned and expressed as a fusion protein with GST, and the expressed soluble fusion protein was purified by affinity chromatography (**Chapter 4**). This procedure yielded very low quantities of active enzyme for biochemical characterization. Using the same expression construction, we decided to proceed with the purification and refolding of the insoluble form of the recombinant GST-babesipain-1, which constituted the bulk of the expressed fusion protein. Insoluble inclusion bodies were washed with a urea buffer and later solubilised with 8M Urea (Fig. 5.1A). Afterwards, the denatured and reduced GST-babesipain-1 protein was refolded in more than 100 buffer testing conditions in a

microplate format. Aggregation suppressors (L-arginine, KCl, Triton X-100, and polyethylene glycol), and co-solvents (glycerol and sucrose) did not influenced positively the refolding. The best tested conditions for the refolding of denatured babesipain-1 to an active enzyme were obtained with basic pH and a GSH:GSSG ratio of 5:1 or 5:0.5 (Table 5.1). The refolded protein samples showed a near to exponential increase in activity assays during the initial time period of analysis, suggesting that upon acidification to pH 5.0, the refolded pro-enzyme processed itself into a mature and more active enzyme.

Table 5.1. Optimization of refolding of babesipain-1. Effect of pH and GSH:GSSG ratio (concentrations in mM) on the efficiency of the refolding of babesipain-1 in 20 mM Tris. Activity of the hydrolysis of Z-Phe-Arg-AMC for each reaction were normalized against a maximum of 100 to express refolding efficiency.

pH	7.5	8.0	8.5	9.0	9.5	10.0	10.5	11.0	11.5	12.0
GSH:GSSG	(Activity, % of maximum)									
10:1	0.0	0.3	35.7	56.5	68.3	90.5	83.2	92.3	83.3	89.7
10:0.5	0.3	0.8	44.8	64.9	74.3	89.6	88.1	98.0	87.5	97.3
5:1	0.1	24.4	54.6	72.5	77.8	91.6	84.1	94.8	89.4	98.6
5:0.5	0.2	30.4	62.7	72.0	86.0	96.7	96.8	100.0	99.0	99.5
2.5:0.5	0.9	31.6	51.5	61.3	72.7	80.1	77.2	87.0	76.7	85.3
1:0.5	3.6	18.7	31.7	39.2	45.8	51.1	49.1	54.7	53.5	57.2
0.5:0.5	2.8	11.6	20.8	25.6	30.5	32.9	34.1	39.1	37.0	39.9
0:0	1.1	3.1	6.1	9.4	13.9	17.5	19.5	20.7	20.3	28.2

Large-scale refolding of babesipain-1 was carried out by 100-fold rapid dilution of the denatured protein in optimized buffer (20 mM Tris, 1 mM EDTA, 5 mM GSH and 0.5 mM GSSG, pH 11.0). The profile of the refolded protein material did not suffered significant alterations in comparison with the denatured proteins as observed by SDS-PAGE analysis (Fig. 5.1).

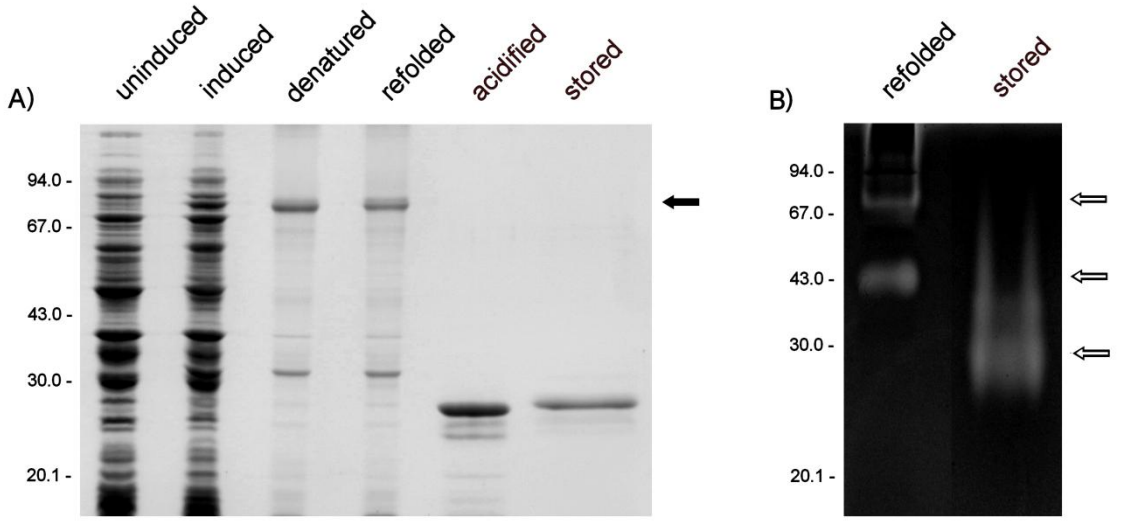

Figure 5.1. Expression, purification and gel substrate activity of recombinant babesipain-1. The expression of the 75 kDa GST-babesipain-1 was evaluated by 12.5 % (w/v) SDS-PAGE (A). Uninduced *E. coli* BL21 cells and induced with IPTG for 3h. Inclusion bodies containing the denatured fusion GST-babesipain-1 protein were purified by Urea washes and later solubilised with 8M Urea. GST-babesipain-1 was refolded, and later acidified to a processed babesipain-1 mixture with GST, and the babesipain-1 mixture with GST was then stored in 50% (v/v) glycerol. Activity was evaluated by gelatin zymography (B) in a one month old refolding sample at pH 9.0 and in 50% glycerol babesipain-1. The positions of molecular weight markers are labelled in kDa. Black arrow indicates the GST-babesipain-1 fusion protein and white arrows indicate activity.

The refolded pro-enzyme was then acidified to promote auto-activation. The conditions for the auto-activation of babesipain-1 were optimized in respect to pH (4.0 to 6.0) and temperature (4 °C to 37 °C). The best pH and temperature was defined as the condition at which the maximum of activity in the acidified samples was obtained after 2h and 4h of incubation. The best pH and temperature for auto-activation were 5.0 and 4 °C. Low temperatures are routinely used to prevent agglomeration. Despite this, and at all temperatures, a significant amount of refolded protein precipitated almost immediately after change in pH, and at lower values than the optimum pH 5.0, the precipitated material increased significantly. Part of the precipitated material is composed by *E. coli* contaminating proteins, as these are not observed in SDS-PAGE analysis after acidification (Fig. 5.1A). Upon acidification, a change in the mobility of the 75 kDa GST-babesipain-1 was observed in the SDS-PAGE analysis (Fig. 5.1A). The auto-activation was confirmed

by gelatinase activity gel, as activity was observed as a colourless band at approximately 30 kDa, the estimated molecular weight of the mature babesipan-1. The auto-activation of babesipain-1 is a multi-step process, as an intermediate product with gelatinase activity at 43 kDa was observed in an old refolding sample (Fig. 5.1B), in which the process of maturation occurred slowly, since the storage conditions of pH 9.0 and 4 °C were not the most favourable. The majority of the protein observed at 27 kDa (Fig. 5.1A) was later confirmed as being the GST as it was removed by affinity purification. Although the removal of GST did not alter the activity of the mature babesipain-1, it changed the stability as observed as a rapid loss of enzyme activity in assays with the peptide Z-Phe-Arg-AMC. Cysteine proteases are usually relatively unstable enzymes which are prone to autocatalytic degradation, oxidation, and protein denaturation (Brömme et al., 2004). The more stable protein mixture of GST and mature babesipain-1 (active enzyme 5.6% of total protein) was then used in all subsequent kinetic experiments.

Babesipain-1 had typical properties of a papain-family cysteine protease, with acidic pH optimum and requirement for a reducing environment for maximum activity (Fig. 5.2A and B). Babesipain-1 showed a broad activity spectrum under the pH conditions used, with more than 30% of maximum activity from pH 4.0 to pH 9.0.

The interaction of babesipain-1 with inhibitors of cysteine proteases and other classes was also analysed (Fig. 5.2C). As is typical of cysteine proteases, babesipain-1 was potently inhibited by E-64, ALLN, leupeptin and cystatin. Inhibitors of other classes of proteases only had a minor effect or no effect.

The ability of the mature form of babesipain-1 to cleave a variety of peptide substrates was examined (Table 5.2). Babesipain-1 preferred substrates with Valine at the P_2 position with a P_2 rank order of Val > Leu > Phe in contrast with falcipain-2 and -3 from *P. falciparum* that preferred Leu > Phe > Val (Sigwali et al., 2001). The higher specificity constant (k_{cat}/K_m) for the substrate with Valine at the P_2 position was a result of increased substrate turnover (kcat), as babesipain-1 showed apparent higher affinity (lower K_m) for the interaction with substrates with Leucine at the P_2 position.

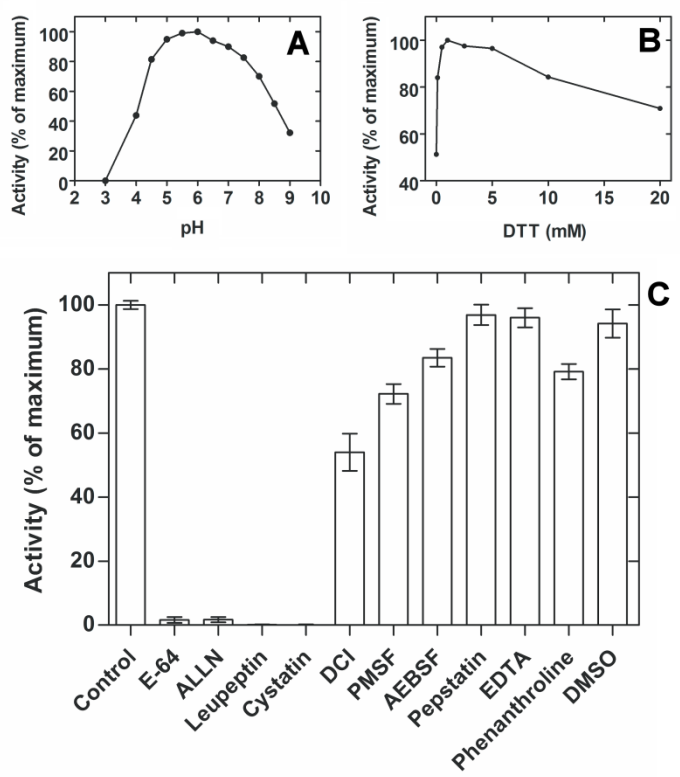

Figure 5.2. Effect of pH, reductant concentration, and inhibitors on the activity of babesipain-1. The activity of recombinant babesipain-1 was assayed against 50 µM Z-Phe-Arg-AMC in 100 mM sodium chloroacetate (pH 3.0), sodium acetate (pH 3.5–5.5), sodium phosphate (pH 6.0–7.0) or Tris (7.5–9.0) and 1 mM DTT (A) and in 100 mM sodium acetate pH 5.5, with different concentrations of DTT (B). For inhibitor assays, recombinant babesipain-1 was added to 50 µM Z-Phe-Arg-AMC in 100 mM sodium acetate pH 5.5, in the absence or presence of E-64 (1 µM), ALLN (100 µM), leupeptin (100 µM), cystatin (10 µg/mL), DCI (50 µM), PMSF (1 mM), AEBSF (1 mM), pepstatin (8 µM), EDTA (1 mM), 1,10-phenanthroline (0.5 mM) or DMSO 1 % (v/v) (C). Release of AMC was continuously monitored for 30 min, slopes of fluorescence over time were calculated, and results were expressed as percentage of maximum activity for each assay.

Table 5.2. Kinetic parameters for substrate hydrolysis by recombinant babesipain-1.

Substrate	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} M^{-1})
Z-Phe-Arg-AMC	59.0	0.306	5.19×10^3
Z-Leu-Arg-AMC	14.6	0.333	2.27×10^4
Z-Leu-Leu-Arg-AMC	10.4	0.245	2.37×10^4
Z-Val-Val-Arg -AMC	22.7	0.681	3.00×10^4

5.4. Discussion

We have characterized for the first time the recombinant babesipain-1, after successfully improved the production of active recombinant babesipain-1 using a different approach in relation to a previous study (**Chapter 4**). Babesipain-1 had typical properties of a papain-family cysteine protease, including hydrolysis of typical papain-family peptide substrates (Table 5.2), an acidic pH optimum, requirement for a reducing environment for maximum activity, and inhibition by standard cysteine protease inhibitors (Fig. 5.2). Two potent inhibitors of babesipain-1, E-64d and ALLN, were also found to effectively inhibit the *in vitro* growth of *B. bovis* (Okubo et al., 2007). Invasion of erythrocytes by *B. bovis* was only inhibited by E-64d and not ALLN (Okubo et al., 2007), suggesting that another cysteine protease may be involved in this role. In the related Apicomplexan parasite, *P. falciparum*, the cysteine proteases falcipain-2 and -3 have been implicated in essential roles as host haemoglobin degradation (Sijwali et al., 2001b), and rupture of infected erythrocytes (Dua et al., 2001; Hanspal et al., 2002). Intracellular processing of these falcipains was inhibited by E-64d and ALLN (Dahl and Rosenthal, 2005). Accordingly, it is possible that babesipain-1 may digest host proteins for metabolism or participate in the rupture of erythrocytes. Although, pseudo food vacuoles are formed by enclosing host cytoplasm, intraerythrocytic *Babesia* spp. do not have haemozoin, the breakdown product of haemoglobin digestion (Kawai et al. 1999a), in contrast with *Plasmodium* spp. that use haemoglobin as a major source of nutrients and have haemozoin (malaria pigment) (Francis et al., 1997). In addition, *Babesia* spp. induce severe ultrastructural changes in the erythrocyte membrane and shape, with tubular structures extending from the surface of the infected erythrocyte membrane directly to the parasites, and it was suggested that *Babesia* spp. come into direct contact with the host plasma and may obtain nutrients from both outside and inside host cells (Braga et al., 2006; Kawai et al., 1999a; Kawai et al., 1999b). It is still unclear the feeding mechanisms of *Babesia* spp. and if it involves cysteine proteases. The availability of recombinant cysteine proteases may aid in clarifying this subject.

At slightly basic pH 7.5, falcipain-2 with around 40% of maximum activity is more active against peptide substrates than falcipain-3 with only 5% (Sijwali et al., 2001b). Accordingly, falcipain-2 cleaves cytoskeletal proteins (pH optimum of 7.5 to 8.0) and may facilitate erythrocyte rupture by mature schizonts (Dua et al., 2001; Hanspal et al., 2002). Babesipain-1 also had considerable activity against a peptide substrate at slightly basic pH 7.5 (>80% maximum activity), although it remains uncertain if this is the result of an adaptation to the absence of an acidic environment as the plasmodial food vacuole, or if babesipain-1 participates in degradation of host proteins in the cytosol where a slightly basic optimum pH would be ideal.

The nature of the S2 pocket and in particular the residue present in the hollow end of the pocket is essential to the substrate specificity of clan CA enzymes. The S2 pocket is usually constituted by hydrophobic residues, but the key residue (residue 205 in papain) present at the bottom is not conserved (Sajid and McKerrow, 2002). Falcipain-2 and -3 have in this critical position the polar residues Asp234 and Glu243, respectively (Kerr et al., 2009). Babesipain-1 has a bulky hydrophobic phenylalanine in the end of the S2 pocket, an unusual feature in common with the human cathepsin W and with the CP1 protease from the cellular slime mold *Dictyostelium discoideum* (Brinkworth et al., 2000). Babesipain-1 showed a P_2 rank order of Val > Leu > Phe in contrast with falcipain-2 and -3 that preferred Leu > Phe > Val (Sigwali et al., 2001). The preference of babesipain-1 for the small residue valine, suggests a narrow S2 Pocket and therefore restricted substrate specificity, although the overall environment is of great importance and not only the nature of the end of the S2 pocket.

In light of recent work on the inhibition of cultures of *B. bovis* parasites (Okubo et al., 2007) suggesting babesipain-1 as a potential drug target, the availability of this recombinant protein can facilitate research in this area.

Chapter 6

General Discussion and Conclusions

6.1. Discussion

Bovine babesiosis is a tick–borne disease with significant morbidity and mortality, that impose severe restrictions to cattle movement, and the economic losses associated to this disease can be considerable (Wright, 1990). Control measures of bovine babesiosis include the eradication or reduction of ticks, diagnosis, vaccination and treatment. The increasing resistance of ticks to acaricides constitutes a current and serious problem in the management of tick control (Graf et al., 2004). Moreover, live attenuated vaccines are used for the control of babesiosis in many parts of the world, but rely on region-specific attenuated strains for which vaccine breakthrough is not uncommon (de Waal and Combrink, 2006). Outbreaks also occur in endemic areas especially after cattle movements and consequent challenge with different strains (Alfredo et al., 2005; Bock et al., 2004). Therefore, control of babesiosis relies in the prompt diagnosis and effective treatment as the last alternative. Detection of infections for epidemiologic purposes is also important to evaluate the effectiveness of adopted control measures. This work was designed to contribute to the improvement of some of the control measures, and the objectives of this study include: the development of a new molecular diagnostic method (Chapter 2), survey on the status of babesiosis in Mozambique (Chapter 2 and 3), identification of proteases as potential drug targets (Chapter 4), and biochemical characterization of a cysteine protease as a new drug target (Chapter 5). This work also had the objective to increase the scientific cooperation between Mozambique and Portugal in this important veterinary area. This study has two apparently distinct sections, in diagnosis and treatment areas, but the two areas are closely related, since treatment begins with proper diagnosis, and epidemiological data is fundamental to access the proper treatment. In an endemic region, the treatment should allow the development of immunity, and in other scenarios, like the US free zone from bovine babesiosis, the treatment should result in the complete clearance of the parasites. In the first section of this study, two molecular detection methods were compared with a new method developed in this work, and the status of bovine babesiosis was evaluated using these methods. And, in the second section, particular interest was given to

the identification and characterization of cysteine proteases that are already drug targets for other protozoan parasites.

6.1.1. Development of a new molecular diagnostic method

Diagnosis of clinical cases of babesiosis in Mozambique was commonly based in the classical method of examination of Giemsa stained blood smears from suspected animals. The development of PCR based techniques, lead to new and more sensitive diagnostic methods. Although the molecular diagnostic methods require more instrumentation, they allow for proper identification and classification of pathogens often present with different but morphological similar parasites and in low parasitaemias (Böse et al. 1995). Several methods have been described for detection and differentiation of bovine babesiosis infections including the nested PCR (Figueroa et al., 1993), reverse line blot (RLB) hybridization (Gubbels et al., 1999), LAMP (Iseki et al., 2007) and real time PCR (Buling et al., 2007).

In recent years the RLB assay, a molecular based technique that allows the simultaneous detection of several haemoparasites, was implemented in Mozambique (Faculty of Veterinary, Eduardo Mondlane University). Nonetheless, in contrast to what was previously assumed, and although *B. bovis* infections were detected, *B. bigemina* infections were not detected by the RLB assay in previous studies and in the initial assays conducted in this project (Chapter 3). It became therefore necessary to use another method or to develop a new one. Both strategies were followed at the same time in order to successfully detect *B. bigemina* natural infections in Mozambique. In Chapter 2, the development of a new molecular diagnostic method for the detection of *B. bigemina* and *B. bovis* in field samples is described.

Primers were chosen for the new method based on the babesial aspartic protease single copy gene present in both genomes of *B. bigemina* and *B. bovis*. Primers were designed to have a high temperature of melting (Tm), approximately 30 bp long, and 40 to 60% GC content, in order to allow the use of an annealing temperature that is several degrees higher than the Tm of the primers. Although primer annealing and amplification is less efficient at this higher temperature, it is much

more specific (Hecker and Roux, 1996). And to further increase the specificity of the PCR we used a hot-start polymerase, to minimize the yield of non-specific products and the formation of the so-called primer dimer (Kermekchiev et al., 2003).

The new seminested hot-start PCR method was assayed in parallel with the nested PCR described by Figueroa et al. (1993). Both methods allowed the detection of *B. bigemina* and *B. bovis* in high frequency in random field samples from one farm located in Maputo province, Mozambique. However, the seminested hot-start method was comparatively more sensitive: allowed detection in the first PCR in opposition to the old method that only permitted detection in the second or nested PCR; and, the new method was more specific in the amplification of *B. bigemina* amplicons than the old method. The combination of long primers and a hot-start polymerase were determinant to the improvement of the specificity of the PCR and consequently to the improvement of the sensitivity of the method. This strategy can be used in the development of other molecular detection methods for the detection of other pathogens, or in the improvement of previous methodologies.

The new method requires validation for use in other regions of the world that may have local strains with sequence differences in the primers regions, although in the development of the method, amplifications from genomic DNA of Israeli strains of *B. bigemina* and *B. bovis* were positive.

The LAMP assay (Iseki et al., 2007) and the real time PCR (Buling et al., 2007) were not used in this study since they were published after the initial development of the seminested hot-start PCR. These methods also require specific reagents and/or proper apparatus, not available in Mozambique where the method is to be implemented.

The seminested hot-start PCR method was then used to study the status of bovine babesiosis in different geographical locations of Maputo province as described in Chapter 3.

6.1.2. Survey on the status of babesiosis in Mozambique

The study on bovine babesiosis in Mozambique was extended with the collection of blood samples in other 4 farms geographically well distributed through Maputo

province. In total, 477 samples from 5 farms were analyzed by two different methods, the seminested hot-start PCR and the RLB assay, with the objective to evaluate the prevalence and distribution of bovine babesiosis and of other haemoparasites.

The detection results differed considerably between both methods and locations. Using the seminested hot-start PCR, detection of *B. bigemina* between farms varied from 30% to 89% with an overall detection of 61%, and detection of *B. bovis* ranged between 27% and 83% with an overall frequency of 53%. Using the RLB assay, *B. bigemina* was not detected and detection of *B. bovis* ranged between 0% and 17% with a total frequency of 5.1%.

Analysis of new sequences of the 18S rRNA gene from *B. bigemina* obtained in this study, revealed that the current *B. bigemina* RLB probe is unspecific for the identification of all the identified isolates from Mozambique, as was also recently reported for Argentinean isolates (Petrigh et al., 2008). This clearly explains the 0% detection of *B. bigemina* by the RLB. Nevertheless, the detection of *B. bovis* by the RLB (17%) was also below the detection results by the seminested hot-start PCR (53%). This difference may result from the probe for *B. bovis* being unspecific to all the isolates or lower sensitivity of the RLB due to PCR competition when using complex field samples. Similar low detection results using the RLB assay were obtained by Oura et al. (2004) in a survey conducted in Uganda, where *B. bigemina* was detected in 2% of random field samples and *B. bovis* was not detected.

Although the RLB assay was less sensitive in the detection of bovine babesiosis, it allowed the detection in the region of Maputo province of ten different species of the four genera *Babesia*, *Theileria*, *Anaplasma* and *Ehrlichia*. Overall, eleven parasite species were detected by the methods used, and the results show that multiple infections with several species of cattle haemoparasites are common in Maputo province.

The sequences of the 18S rRNA genes cloned in this study also showed that there is microheterogeneity, where minor nucleotide sequence differences occur in homologous regions of the gene (Chae et al., 1999), in this highly conserved gene

between isolates of the same species of *Babesia* or *Theileria*, indicating the high variability within the species in the field.

The variability within species and the presence of common multiple infections that this study showed adds to the presence of several vectors and game reservoirs, and a greater range of susceptibility of bovine breeds, in the complex scenario of host, vector and bovine babesiosis interactions in Mozambique. Accordingly, detection of bovine babesiosis varied greatly between locations. If the higher prevalence in Umbelúzi may be explained by the presence of a susceptible breed in comparison with the remaining locations that have resistant African breeds, the results obtained in the remaining farms may be harder to elucidate. A combination of several factors as different microclimates, disease management, cross treatment with other diseases management (in Bela Vista with trypanosomiasis), and tick species distribution have influence and contribute for the complexity of the management of this disease in Mozambique.

One approach to the control of babesiosis and other TBD rely on the development of endemic stability, where the majority of the cattle population is immune to the disease due to early natural exposure and the development of innate immunity (Mahoney, 1974). The results of this study show that bovine babesiosis is endemic in the Maputo province, although there are some locations with low prevalence numbers that may indicate endemic instability. Cattle owners may opt for intensive tick control, attempting to break the transmission cycle completely. This is much more costly than vaccination, and is problematic if tick control breaks down. In endemic unstable situations, a dry season or intensive tick control can result in drastic reduction of ticks and reduction of transmission to levels at which the disease is not maintained, to create a generation of susceptible cattle that may result in catastrophic outbreaks when are later challenged (Alfredo et al., 2005; Bock et al., 2004). The integration of strategic use of acaricides, the application of vaccines in endemically unstable conditions and the use of tick resistant breeds may in the long term achieve an effective endemic stability (Alfredo et al., 2005), and the availability of a sensitive detection method to access the implementation of these measures is of great importance.

The seminested hot-start PCR was comparatively more sensitive to the RLB assay (Chapter 3) and to the nested PCR (Chapter 2), using random field samples from Mozambique. The seminested hot-start PCR is therefore the more suitable method of the ones tested to use in the future studies on bovine babesiosis in this country.

6.1.3. Identification of proteases as potential drug targets

Proteases found in viruses, bacteria, and parasites, are recognized as one of the largest potential drug target enzyme families, due to their lower homology to their mammalian orthologs, and offer target opportunities to identify selective inhibitors that have minimal cross-reactivity with host proteases (McKerrow et al., 2008).

In chapter 4, the data-mining of the genome of *B. bigemina* for the identification of proteases as potential drug targets is described. The focus was on the cysteine class of proteases that are known to be essential to the life cycle and pathogenicity of parasitic protozoa (Sajid and McKerrow, 2002). Cysteine proteases have numerous roles in Apicomplexa parasites such as catabolic functions, immunoevasion, enzyme activation, cellular invasion and rupture (Rosenthal, 2004; Dowse and Soldati, 2004). Hence, cysteine proteases of protozoan parasites are recognized drug targets and specific inhibitors are in validation for chemotherapy of leishmaniasis, malaria, and trypanosomiasis (McKerrow et al., 2008).

Five genes were identified by sequence similarity search to be homologous to the cysteine protease family in the ongoing *B. bigemina* genome sequencing project database and were compared with the annotated genes from the complete bovine piroplasms genomes of *B. bovis*, *Theileria annulata*, and *T. parva*. Multiple genome alignments and sequence analysis were used to evaluate the molecular evolution events that occurred in the C1 family of cysteine proteases in these piroplasms of veterinary importance.

Analysis and comparisons between the genomes of bovine piroplasms revealed that there are five distinct groups of cysteine protease genes of Family C1 in *B. bigemina*, four groups in *B. bovis* and six groups in *Theileria* spp.. All analysed species have two cathepsin C-like groups, and therefore the differences observed

are in the number of cathepsin L-like groups. The overall content of cysteine protease genes suggests that after the separation from a common ancestor, *Babesia* lost cysteine protease genes (deletion of the SERA gene, also present in *Plasmodium* spp.) and *Theileria* gained through duplication of the existent ones (13 cathepsin L and C-like genes present in both *T. annulata* and *T. parva* against only 5 in *B. bigemina* and 4 in *B. bovis*). Molecular evolution in *Theileria* occurred not only in a sheer increase in the number of genes, but also in sequence diversity.

The active sites of the cathepsin L-like cysteine proteases of *Theileria* have unusual mutations and insertions that show no parallelism in *Babesia* orthologs that have conserved catalytic residues and hydrophobic residues in the S2 pocket responsible for the substrate preference to hydrophobic residues in the S2 subsite, characteristic of mammalian cathepsin L-like cysteine proteases.

The total differences in number and sequence diversity observed in the cathepsin L-like cysteine proteases of these Genera may in part explain the rather different mechanisms of infection in the host vertebrates. *B. bigemina* and *B. bovis* have a low number of cathepsin L-like cysteine proteases (3 and 2 genes respectively) in comparison with the closely related *Theileria* spp. (11 genes) and also with *Plasmodium* spp. (16 genes in *P. falciparum*)(Wu et al., 2003; Rosenthal, 2004). *Babesia* parasites do not invade lymphocytes in opposition to *Theileria* parasites that have to invade first the lymphocytes in the vertebrate host, and interestingly *Plasmodium* parasites also invade hepatocytes before invading erythrocytes. In result of this, *Theileria* and *Plasmodium* need to invade and promote rupture of more types of cells in comparison to *Babesia*, and consequently have more and different substrate targets that require different and specific enzymatic activities. Since cysteine proteases are involved in invasion or egress in other parasitic protozoa including *Plasmodium* spp. (Dowse and Soldati, 2004; Rosenthal, 2004; Sajid and McKerrow, 2002; Yeoh et al., 2007) the divergent cysteine proteases of *Theileria* probably have roles in the invasion or egress in the lymphocyte stage, and this may explain the observed evolution through duplication and diversity of these enzymes.

One of the three cathepsin L-like enzymes of *B. bigemina*, babesipain-1, was expressed in a recombinant expression system, and for the first time the activity of

a cysteine protease from piroplasms was described. The activity of babesipain-1 was completely inhibited by E-64, and therefore, babesipain-1 is one of the possible targets of the inhibitor E-64d that effectively inhibited *in vitro* the growth of *B. bovis* parasites and invasion of host erythrocytes (Okubo et al., 2007), and indicated that cysteine proteases are essential to the life cycle of *Babesia*. The reduced number of cathepsin L-like cysteine proteases from *Babesia* decreases the possibilities of occurrence of functional redundancy, where different enzymes have similar importance in a specific role, and increases the probability of development of inhibitors that block an essential role and thus prevents the development of resistance. Therefore, babesipain-1 and the other two babesial cathepsin L-like cysteine proteases retain their importance as potential drug targets.

6.1.4. Biochemical characterization of babesipain-1 as a new drug target

Babesipain-1 was expressed as a fusion protein with gluthatione S-transferase (GST) and the soluble protein purified by affinity chromatography. The recombinant babesipain-1 showed activity against typical peptide substrates of cysteine proteases, and was inhibited by E-64, but the low yield of the soluble purification prevented additional characterization (Chapter 4). Using the same expression construction, we proceeded with the purification and refolding of the insoluble form of the recombinant GST-babesipain-1, which constituted the bulk of the expressed fusion protein. Conditions for the refolding and activation of the pro-enzyme to a mature form were successfully optimized to yield sufficient amounts of enzyme to proceed with the biochemical characterization of babesipain-1 (Chapter 5).

Babesipain-1 showed typical properties of a papain-family cysteine protease, including hydrolysis of typical papain-family peptide substrate, an acidic pH optimum, requirement for a reducing environment for maximum activity, and inhibition by standard cysteine protease inhibitors.

The optimum pH for the protease activity against peptide substrates was 5.5, but enzymatic activity was observed between pH 4.0 and pH 9.0, a wider range than the observed for the cysteine proteases falcipain-2 and -3 from *P. falciparum*, pH 4.5 to 8.0 and pH 4.0 to 7.5 respectively (Sijwali et al., 2001b). At slightly basic pH

7.5, babesipain-1 maintained 83% of maximum activity, falcipain-2 less than 50% and falcipain-3 showed no activity (Sijwali et al., 2001b). This suggests a role for babesipain-1 in the degradation of host proteins in the cytosol where a slightly basic optimum pH would be ideal. Falcipain-2 cleaves cytoskeletal proteins, at a pH optimum of 7.5 to 8.0, and facilitates erythrocyte rupture by mature schizonts (Dua et al., 2001; Hanspal et al., 2002), and babesipain-1 may play a similar role.

6.2. Conclusions

A new molecular method for the direct detection of *B. bovis* and *B. bigemina* was developed for the analysis of samples from Mozambique and the new method was more sensitive than other two widely used methods.

Bovine babesiosis in Mozambique is widely distributed and common in the province of Maputo, but there are locations with low prevalence of this disease, and the desired status of endemic stability cannot be therefore guaranteed.

Cysteine proteases from *Babesia* spp. were identified and are promising drug targets, due to their reduced numbers in the genomes and essential roles in the life cycle. The recombinant and active cysteine protease babesipain-1 was expressed and characterized and its availability allows now the screening for specific inhibitors.

6.3. Future work

The new seminested hot-start PCR can now be used in epidemiological studies in Mozambique to address the relationship between babesiosis and the factors affecting it´s transmission: vector, climate, tick control, and bovine breed resistance.

The seminested hot-start PCR should be validated for global use, by testing samples from different parts of the world and comparing the results with another method results to detect the existence of different strains that may not be detected by the current primers. Alternatively, the design of primers against a multi copy gene instead of the single copy aspartic protease gene, in combination with the

long primers and hot-start polymerase strategy may increase the sensitivity in the first PCR, excluding the need of a second PCR and the contamination issues associated.

The importance of the different babesipains in the infection process in the vertebrate host and consequently as drug targets can be studied by first evaluating the expression profile of these cysteine protease genes in different life cycle stages using RT-PCR, and second by gene disruption using RNA interference or transgenetics, although both techniques have to be first developed for *Babesia* species.

A different expression and purification method for babesipain-1 should be attempted since the enzyme was obtained in a final mixture of partially processed pro-enzyme, GST and mature enzyme, and this crude mixture may interfere in the final activity profile. A truncated form of the gene can be expressed in a recombinant system without the majority of the pro-part and with a smaller tag like poly histidine, which allows the purification of the insoluble form in opposition to GST that only allows the purification of the soluble form.

Recombinant active babesipain-1 can be used in the screening for specific inhibitors in a microtiter format for large scale analysis. The selected inhibitors can then be evaluated using *in vitro* cultures.

References

Alfredo, A.A.N., Jonsson, N.N., Finch, T.M., Neves, L., Molloy, J.B., Jorgensen, W.K., 2005. Serological survey of *Babesia bovis* and *Anaplasma marginale* in cattle in Tete Province, Mozambique. Trop. Anim. Health Pro. 37, 121-131.

Allsopp, M.T., Cavalier-Smith, T., de Waal, D.T., Allsopp, B.A., 1994. Phylogeny and evolution of the piroplasms. Parasitol. 108, 147–152.

Almeria, S., Castellà, J., Ferrer, D., Ortuño, A., Estrada-Peña, A., Gutiérrez, J.F., 2001. Bovine piroplasms in Minorca (Balearic Islands, Spain): a comparison of PCR-based and light microscopy detection. Vet. Parasitol. 99, 249-259.

Aly, A.S., Matuschewski, K., 2005. A malarial cysteine protease is necessary for *Plasmodium* sporozoite egress from oocysts. J. Exp. Med. 202, 225-230.

Arisue, N., Hirai, M., Arai, M., Matsuoka, H., Horii, T., 2007. Phylogeny and evolution of the SERA multigene family in the genus *Plasmodium*. J. Mol. Evol. 65, 82-91.

Babes, V., 1888. Sur l'hemoglobinurie bacterienne du boeuf. Comptes rendus hebdomadaires des seances de l'Academie des Sciences, Paris 107, 692–694.

Barrett, A.J., Kirschke, H., 1981. Cathepsin B, Cathepsin H, and Cathepsin L. Methods Enzymol. 80, 535-561.

Baylis, H.A., Megson, A., Mottram, J.C., Hall, R., 1992. Characterisation of a gene for a cysteine protease from *Theileria annulata*. Mol. Biochem. Parasitol. 54, 105-107.

Bekker, C.P., de Vos, S., Taoufik, A., Sparagano, O.A., Jongejan, F., 2002. Simultaneous detection of *Anaplasma* and *Ehrlichia* species in ruminants and

detection of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks by reverse line blot hybridization. Vet. Microbiol. 89, 223-238.

Binder, E.M., Kim, K., 2004. Location, Location, Location: Trafficking and Function of Secreted Proteases of *Toxoplasma* and *Plasmodium*. Traffic 5, 914-924.

Birkenheuer, A.J., Levy, M.G., Breitschwerdt, E.B., 2003. Development and Evaluation of a Seminested PCR for Detection and Differentiation of *Babesia gibsoni* (Asian Genotype) and *B. canis* DNA in Canine Blood Samples. J. Clin. Microbiol. 41, 4172-4177.

Bock, R., Jackson, L., de Vos, A., Jorgensen, W., 2004. Babesiosis of cattle. Parasitol. 129, S247-S269.

Bock, R.E., De Vos, A.J., Kingston, T.G., Mclellan, D.J., 1997. Effect of breed of cattle on innate resistance to infection with *Babesia bovis*, *Babesia bigemina* and *Anaplasma marginale*. Aust. Vet. J. 75, 337-40.

Braga, W., Venasco, J., Willard, L., Moro, M.H., 2006. Ultrastructure of *Babesia* WA1 (Apicomplexa: Piroplasma) during infection of erythrocytes in a hamster model. J. Parasitol. 92, 1104-1107.

Brayton, K.A., Lau, A.O., Herndon, D.R., Hannick, L., Kappmeyer, L.S., Berens, S.J., Bidwell, S.L., Brown, W.C., Crabtree, J., Fadrosh, D., Feldblum, T., Forberger, H.A., Haas, B.J., Howell, J.M., Khouri, H., Koo, H., Mann, D.J., Norimine, J., Paulsen, I.T., Radune, D., Ren, Q., Smith Jr., R.K., Suarez, C.E., White, O., Wortman, J.R., Knowles Jr., D.P., McElwain, T.F., Nene, V.M., 2007. Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. PLoS Pathog. 3, 1401-1413.

Brígido, C., da Fonseca, I.P., Parreira, R., Fazendeiro, I., do Rosário, V.E., Centeno-Lima, S., 2004. Molecular and phylogenetic characterization of *Theileria*

spp. parasites in autochthonous bovines (Mirandesa breed) in Portugal. Vet. Parasitol. 123, 17-23.

Brinkworth, R.I., Tort, J.F., Brindley, P.J., Dalton, J.P., 2000. Phylogenetic relationships and theoretical model of human cathepsin W (lymphopain), a cysteine proteinase from cytotoxic T lymphocytes. Int. J. Biochem. Cell Biol. 32, 373-384.

Brömme, D., Nallaseth, F.S., Turk, B., 2004. Production and activation of recombinant papain-like cysteine proteases. Methods 32, 199-206.

Brudno, M., Do, C.B., Cooper, G.M., Kim, M.F., Davydov, E., Green, E.D., Sidow, A., Batzoglou, S., 2003. LAGAN and Multi-LAGAN: efficient tools for largescale multiple alignment of genomic DNA. Genome Res. 13, 721-731.

Buling, A., Criado-Fornelio, A., Asenzo, G., Benitez, D., Barba-Carretero, J.C., Florin-Christensen, M., 2007. A quantitative PCR assay for the detection and quantification of *Babesia bovis* and *B. bigemina*. Vet. Parasitol. 147, 16-25.

Böse, R., Jorgensen, W.K., Dalgliesh, R.J., Friedhoff, K.T., De Vos, A.J., 1995. Current state and future trends in the diagnosis of babesiosis. Vet. Parasitol. 57, 61–74.

Calder, J.A.M., Reddy, G.R., Chieves, L., Courtney, C.H., Littell, R., Livengood, J.R., Norval, R.A.I., Smith, C., Dame, J.B., 1996. Monitoring *Babesia bovis* infections in cattle by using PCR-based tests. J. Clin. Microbiol. 34, 2748-2755.

Campanella, J.J., Bitincka, L., Smalley, J., 2003. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. BMC Bioinf. 4, 29.

Cavalier-Smith, T., 1993. Kingdom Protozoa and Its 18 Phyla. Microbiol. Rev. 57, 953-994.

Centeno-Lima, S., do Rosário, V., Parreira, R., Maia, A.J., Freudenthal, A.M., Nijhof, A.M., Jongejan, F., 2003. A fatal case of human babesiosis in Portugal: molecular and phylogenetic analysis. Trop. Med. Int. Health 8, 760-764.

Chae, J.S., Allsopp, B.A., Waghela, S.D., Park, J.H., Kakuda, T., Sugimoto, C., Allsopp, M.T., Wagner, G.G., Holman, P.J., 1999. A study of the systematics of *Theileria* spp. based upon small-subunit ribosomal RNA gene sequences. Parasitol. Res. 85, 877-883.

Costa-Júnior, L.M., Rabelo, E.M.L., Martins Filho, O.A., Ribeiro, M.F.B., 2006. Comparison of different direct diagnostic methods to identify *Babesia bovis* and *Babesia bigemina* in animals vaccinated with live attenuated parasites. Vet. Parasitol. 139, 231-236.

Dahl, E.L., Rosenthal, P.J., 2005. Biosynthesis, localization, and processing of falcipain cysteine proteases of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 139, 205-212.

Dasaradhi, P.V., Mohmmed, A., Kumar, A., Hossain, M.J., Bhatnagar, R.K., Chauhan, V.S., Malhotra, P., 2005. A role of falcipain-2, principal cysteine proteases of *Plasmodium falciparum* in merozoite egression. Biochem. Biophys. Res. Commun. 336, 1062-1068.

de Vos, A.J., 1979. Epidemiology and control of bovine babesiosis in South Africa. J. S. Afr. Vet. Assoc. 50, 357–362.

de Vos, A.J., Dalgliesh, R.J., Mcgregor, W., 1986. Effect of imidocarb dipropionate prophylaxis on the infectivity and immunogenicity of a *Babesia bovis* vaccine in cattle. Aust. Vet. J. 63, 174–178.

de Vries, E., Corton, C., Harris, B., Cornelissen, A.W.C.A., Berriman, M., 2006. Expressed sequence tag (EST) analysis of the erythrocytic stages of *Babesia bovis*. Vet. Parasitol. 138, 61-74.

de Waal D.T., Combrink, M.P., 2006. Live vaccines against bovine babesiosis. Vet. Parasitol. 138, 88-96.

Dipeolu, O.O., Amoo, A., 1984. The presence of kinetes of a *Babesia* species in the haemolymph smears of engorged hyalomma ticks in Nigeria. Vet. Parasitol. 17, 41-46.

Dowse, T., Soldati, D., 2004. Host cell invasion by the apicomplexans: the significance of microneme protein proteolysis. Curr. Opin. Microbiol. 7, 388-396.

Doyle, P.S., Zhou, Y.M., Engel, J.C., McKerrow, J.H., 2007. A cysteine protease inhibitor cures Chagas' disease in an immunodeficient-mouse model of infection. Antimicrob. Agents Chemother. 51, 3932-3939.

Dua, M., Raphael, P., Sijwali, P.S., Rosenthal, P.J., Hanspal, M., 2001. Recombinant falcipain-2 cleaves erythrocyte membrane ankyrin and protein 4.1. Mol. Biochem. Parasitol. 116, 95-99.

Dunn, L.A., Andrews, K.T., McCarthy, J.S., Wright, J.M., Skinner-Adams, T.S., Upcroft, P., Upcroft, J.A., 2007. The activity of protease inhibitors against *Giardia duodenalis* and metronidazole-resistant *Trichomonas vaginalis*. Int. J. Antimicrob. Agents. 29, 98-102.

Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39, 783-791.

Fetterer, R.H., Miska, K.B., Lillehoj, H., Barfield, R.C., 2007. Serine protease activity in developmental stages of *Eimeria tenella*. J. Parasitol. 93, 333-340.

Figueroa, J.V., Chieves, L.P., Johnson, G.S., Buening, G.M., 1993. Multiplex polymerase chain reaction based assay for the detection of *Babesia bigemina*, *Babesia bovis* and *Anaplasma marginale* DNA in bovine blood. Vet. Parasitol. 50, 69-81.

Fish, L., Leibovich, B., Krigel, Y., McElwain, T., Shkap, V., 2008. Vaccination of cattle against *B. bovis* infection with live attenuated parasites and non-viable immunogens. Vaccine 26S, G29-G33.

Francis, S.E., Sullivan Jr., D.J., Goldberg, D.E., 1997. Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. Annu. Rev. Microbiol. 51, 97-123.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., Dubchak, I., 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32, W273-279.

Gardner, M.J., Bishop, R., Shah, T., de Villiers, E., Canton, J.M., Hall, N., Ren, Q., Paulsen, I.T., Pain, A., Berriman, M., Wilson, R.J.M., Sato, S., Stuart, A.R., Mann, D.J., Xiong, Z., Shallom, S.J., Weidman, J., Jiang, L., Lynn, J., Weaver, B., Shoaibi, A., Domingo, A.R., Wasawo, D., Crabtree, J., Wortman, J.R., Haas, B., Angiuoli, S.V., Creasy, T.H., Lu, C., Joanna-Silva, B., Utterback, T.R., Feldblyum, T.V., Pertea, M., Allen, J., Nierman, W.C., Taracha, E., Salzberg, S.L., White, O.R., Fitzhugh, H.A., Morzaria, S., Venter, J.C., Fraser, C.M., Nene, V., 2005. Genome sequence of *Theileria parva*, a bovine pathogen causing a lymphoproliferative disease. Science 309, 134-137.

Gayo, V., Romito, M., Nel, L.H., Solari, M.A., Viljoen, G.J., 2003. PCR-based detection of the transovarial transmission of Uruguayan *Babesia bovis* and *Babesia bigemina* vaccine strains. Onderstepoort J. Vet. Res. 70, 197-204.

Georges, K., Loria, G.R., Riili, S., Greco, A., Caracappa, S. Jongejan, F., Sparagano, O., 2001. Detection of haemoparasites in cattle by reverse line blot

hybridisation with a note on the distribution of ticks in Sicily. Vet. Parasitol. 99, 273-286.

Giegerich, R., Meyer, F., Schleiermacher, C., 1996. GeneFisher - software support for the detection of postulated genes. Proc. Int. Conf. Intell. Syst. Mol. Biol. 4, 68-77.

Goff, W.L., Johnson, W.C., Molloy, J.B., Jorgensen, W.K., Waldron, S.J., Figueroa, J.V., Matthee, O., Adams, D.S., McGuire, T.C., Pino, I., Mosqueda, J., Palmer, G.H., Suarez, C.E., Knowles, D.P., McElwain, T.F., 2008. Validation of a competitive enzyme-linked immunosorbent assay for detection of *Babesia bigemina* antibodies in cattle. Clin. Vaccine Immunol. 15, 1316-1321.

Goff, W.L., Molloy, J.B., Johnson, W.C., Suarez, C.E., Pino, I., Rhalem, A., Sahibi, H., Ceci, L., Carelli, G., Adams, D.S., McGuire, T.C., Knowles, D.P., McElwain, T.F., 2006. Validation of a competitive enzyme-linked immunosorbent assay for detection of antibodies against *Babesia bovis*. Clin. Vaccine Immunol. 13, 1212-1216.

Gough, J.M., Jorgensen, W.K., Kemp, D.H., 1998. Development of tick gut forms of *Babesia bigemina* in vitro. J. Eukaryot. Microbiol. 45, 298-306.

Graf , J.F., Gogolewski, R., Leach-Bing, N., Sabatini, G.A., Molento, M.B., Bordin, E.L., Arantes, G.J., 2004. Tick control: an industry point of view. Parasitol. 129, Suppl., S427-S442.

Gray, J.S., de Vos, A.J., 1981. Studies on a bovine *Babesia* transmitted by *Hyalomma marginatum rufipes* Koch, 1844. Onderstepoort J. Vet. Res. 48, 215-223.

Gubbels, J.M., de Vos, A.P., van der Weide, M., Viseras, J., Schouls, L.M., de Vries, E. Jongejan, F., 1999. Simultaneous detection of bovine *Theileria* and *Babesia* species by reverse line blot hybridization. J. Clin. Microbiol. 37, 1782-1789.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41, 95-98.

Hanspal, M., Dua, M., Takakuwa, Y., Chishti, A.H., Mieuno, A., 2002. *Plasmodium falciparum* cysteine protease falcipain-2 cleaves erythrocyte membrane skeletal proteins at late stages of parasite development. Blood 100, 1048-1054.

He, W., Ohashi, K., Sugimoto, C., Onuma, M., 2005. *Theileria orientalis*: cloning a cDNA encoding a protein similar to thiol protease with haemoglobin-binding activity. Exp. Parasitol. 111, 143-153.

Hecker, K.H., Roux, K.H., 1996 High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. BioTechniques 20, 478–485.

Heussen, C., Dowdle, E.B., 1980. Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. Anal. Biochem. 102, 196-202.

Hildebrandt, A., Tenter, A.M., Straube, E., Hunfeld, K.P., 2008. Human babesiosis in Germany: Just overlooked or truly new? Int. J. Med. Microbiol. 298 (S1), 336–346.

Hilpertshauser, H., Deplazes, P., Schnyder, M., Gern, L., Mathis, A., 2006. *Babesia* spp. Identified by PCR in Ticks Collected from Domestic and Wild Ruminants in Southern Switzerland. Appl. Environ. Microbiol. 72, 6503-6507.

Hirose, M., Akuta, T., Takahashi, N., 1989. Renaturation of ovotransferrin under two-step conditions allowing primary folding of the fully reduced form and the subsequent regeneration of the intramolecular disulfides. J. Biol. Chem. 264, 16867-16872.

Hodder, A.N., Drew, D.R., Epa, V.C., Delorenzi, M., Bourgon, R., Miller, S.K., Moritz, R.L., Frecklington, D.F., Simpson, R.J., Speed, T.P., Pike, R.N., Crabb, B.S., 2003. Enzymic, phylogenetic, and structural characterization of the unusual papain-like protease domain of *Plasmodium falciparum* SERA5. J. Biol. Chem. 278, 48169-48177.

Holman, P.J., Hsieh, M.M., Nix, J.L., Bendele, K.G., Wagner, G.G., Ball, J.M., 2002. A cathepsin L-like cysteine protease is conserved among *Babesia equi* isolates. Mol. Biochem. Parasitol. 119, 295-300.

Homer, M.J., Aguilar-Delfin, I., Telford, S.R., Krause, P.J., Persing, D.H., 2000. Babesiosis. Clin. Microbiol. Rev. 13, 451-469.

Hoyte, H.M., 1961. Initial development of infectious *Babesia bigemina*. Aust. Vet. J. 8, 462–466.

Iseki, H., Alhassan, A., Ohta, N., Thekisoe, O.M.M., Yokoyama, N., Inoue, N., Nambota, A., Yasuda, J., Igarashi, I., 2007. Development of a multiplex loop-mediated isothermal amplification (mLAMP) method for the simultaneous detection of bovine *Babesia* parasites. J. Microbiol. Methods 71, 281–287.

Jenkins, M.C., 2001. Advances and prospects for subunit vaccines against protozoa of veterinary importance. Vet. Parasitol. 101, 291–310.

Kawai, S., Igarashi, I., Abgaandorjiin, A., Miyazawa, K., Ikadai, H., Nagasawa, H., Fujisaki, K., Mikami, T., Suzuki, N., Matsuda, H., 1999. Ultrastructural characteristics of *Babesia caballi* in equine erythrocytes in vitro. Parasitol. Res. 85, 794-799.

Kawai, S., Igarashi, I., Avarzed, A., Ikadai, H., Omata, Y., Saito, A., Nagasawa, H., Toyoda, Y., Suzuki, N., Matsuda, H., 1999b. Tubular structures associated with *Babesia caballi* in equine erythrocytes in vitro. Parasitol. Res. 85, 171-175

Kawai, S., Takahashi, K., Sonoda, M., Kurosawa, T., 1986. Ultra-structure of intra-erythrocytic stages of *Babesia ovata*. Jpn. J. Vet. Sci. 48, 943-949.

Kermekchiev, M.B., Tzekov, A., Barnes, W.M., 2003. Cold-sensitive mutants of Taq DNA polymerase provide a hot-start for PCR. Nucleic Acids Res. 31, 6139-6147.

Kerr, I.D., Lee, J.H., Pandey, K.C., Harrison, A., Sajid, M., Rosenthal, P.J., Brinen, L.S., 2009. Structures of falcipain-2 and falcipain-3 bound to small molecule inhibitors: implications for substrate specificity. J. Med. Chem. 52, 852-857.

Koussis, K., Withers-Martinez, C., Yeoh, S., Child, M., Hackett, F., Knuepfer, E., Juliano, L., Woehlbier, U., Bujard, H., Blackman, M.J., 2009. A multifunctional serine protease primes the malaria parasite for red blood cell invasion. EMBO J. 28, 725-735.

Krishnamurthy, G., Vikram, R., Singh, S.B., Patel, N., Agarwal, S., Mukhopadhyay, G., Basu, S.K., Mukhopadhyay, A., 2005. Hemoglobin receptor in *Leishmania* is a hexokinase located in the flagellar pocket. J. Biol. Chem. 280, 5884-5891.

Lennox, E.S., 1955. Transduction of linked genetic characters of the host by bacteriophage P1. Virology 1, 190-206.

Levine, N.D., 1971. Taxonomy of the piroplasms. Trans. Am. Microsc. Soc. 90, 2–33.

Levine, N.D., 1985. Veterinary Protozoology. Ames, Iowa State University Press.

Liotta, L.A., Stetler-Stevenson, W.G., 1990. Metalloproteinases and cancer invasion. Semin. Cancer Biol. 1, 99-106.

Lüthi, U., 2002. Proteolytic Enzymes as Therapeutic Targets. Eur. Biopharm. Rev. Summer Edition, 2002, 88-93.

Mackenstedt, U., Gauer, M., Fuchs, P., Zapf, F., Schein, E., Mehlhorn, H., 1995. DNA measurements reveal differences in the life cycles of *Babesia bigemina* and *B. canis*, two typical members of the genus *Babesia*. Parasitol. Res. 81, 595–604.

Mahoney, D.F., 1974. The application of epizootiological principals in the control of Babesiosis in cattle. Bulletin of Office International Des Epizooties 81, 123–138.

Mahoney, D.F., Mirre, G.B., 1979. A note on the transmission of *Babesia bovis* (syn. *B. argentina*) by the one-host tick, *Boophilus microplus*. Res. Vet. Sci. 26, 253–254.

Mamman, M., Aliu, Y.O., Peregrine, A.S., 1993. Comparative pharmacokinetics of diminazene in noninfected Boran (*Bos indicus*) cattle and Boran cattle infected with *Trypanosoma* congolense. Antimicrob. Agents Chemother. 37, 1050-1055.

Martins, T.M., Domingos, A., Berry, C., Wyatt, D.M., 2005. The Activity and Inhibition of the Food Vacuole Plasmepsin from the Rodent Malaria Parasite *Plasmodium chabaudi*. Acta Trop. 97, 212-218.

McCoubrie, J.E., Miller, S.K., Sargeant, T., Good, R.T., Hodder, A.N., Speed, T.P., de Koning-Ward, T.F., Crabb, B.S., 2007. Evidence for a common role for the serine-type *Plasmodium falciparum* serine repeat antigen proteases: implications for vaccine and drug design. Infect. Immun. 75, 5565-5574.

McKerrow, J.H., Rosenthal, P.J., Swenerton, R., Doyle, P., 2008. Development of protease inhibitors for protozoan infections. Curr. Opin. Infect. Dis. 21, 668-672.

Mehlhorn, H., Shein, E., 1984. The piroplasms: life cycle and sexual stages. Adv. Parasitol. 23, 37-103.

Molano, M., Meijer, C.J.L.M., Morré, S.A., Pol, R., van den Brule, A.J.C., 2004. Combination of PCR Targeting the VD2 of *omp1* and Reverse Line Blot Analysis for Typing of Urogenital *Chlamydia trachomatis* Serovars in Cervical Scrape Specimens. J. Clin. Microbiol. 42, 2935-2939.

Montero, E., Gonzalez, L.M., Rodriguez, M., Oksov, Y., Blackman, M.J., Lobo, C.A., 2006. A conserved subtilisin protease identified in *Babesia divergens* merozoites. J. Biol. Chem. 281, 35717-35726.

Montero, E., Rafiq, S., Heck, S., Lobo, C.A., 2007. Inhibition of human erythrocyte invasion by *Babesia divergens* using serine protease inhibitors. Mol. Biochem. Parasitol. 153, 80-84.

Mosqueda, J., Falcon, A., Alvarez, J.A., Ramos, J.A., Oropeza-Hernandez, L.F., Figueroa, J.V., 2004. *Babesia bigemina* sexual stages are induced in vitro and are specifically recognized by antibodies in the midgut of infected *Boophilus microplus* ticks. Int. J. Parasitol. 34, 1229-1236.

Mottram, J.C., Frame, M.J., Brooks, D.R., Tetley, L., Hutchison, J.E., Souza, A.E., Coombs, G.H., 1997. The multiple cpb cysteine proteinase genes of *Leishmania mexicana* encode isoenzymes that differ in their stage regulation and substrate preferences. J. Biol. Chem. 272, 14285-14293.

Mukhebi, A.W., Perry, B.D., Kruska, R., 1992. Estimated economics of theileriosis in Africa. Prev. Vet. Med. 12, 73-85.

Nene, V., Gobright, E., Musoke, A.J., Lonsdale-Eccles, J.D., 1990. A single exon codes for the enzyme domain of a protozoan cysteine protease. J. Biol. Chem. 265, 18047-18050.

Nindl, I., Köhler, A., Gottschling, M., Forschner, T., Lehmanna, M., Meijer, C.J.L.M., Snijders, P.J.F., Stockfleth, E., 2007. Extension of the typing in a general-

primer-PCR reverse-line-blotting system to detect all 25 cutaneous beta human papillomaviruses. J. Virol. Meth. 146, 1-4.

North, M.J., Mottram, J.C., Coombs, G.H., 1990. Cysteine proteinases of parasitic protozoa. Parasitol. Today 6, 270-275.

Okubo, K., Yokoyama, N., Govind, Y., Alhassan, A., Igarashi, I., 2007. *Babesia bovis*: effects of cysteine protease inhibitors on in vitro growth. Exp. Parasitol. 117, 214-217.

Oliveira, M.C.S., Oliveira-Sequeira, T.C.G., Araújo, Jr., J.P., Amarante, A.F.T., Oliveira, H.N., 2005. *Babesia* spp. infection in *Boophilus microplus* engorged females and eggs in Sao Paulo State, Brazil. Vet. Parasitol. 130, 61-67.

Oliveira-Sequeira, T.C.G., Oliveira, M.C.S., Araújo, Jr., J.P., Amarante, A.F.T., 2005. PCR-based detection of *Babesia bovis* and *Babesia bigemina* in their natural host *Boophilus microplus* and cattle. Int. J. Parasitol. 35, 105-111.

Oura, C.A.L., Bishop, R.P., Wampande, E.M., Lubega, G.W., Tait, A., 2004. Application of a reverse line blot assay to the study of haemoparasites in cattle in Uganda. Int. J. Parasitol. 34, 603-613.

Pain, A., Renauld, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W., Kerhornou, A., Aslett, M., Bishop, R., Bouchier, C., Cochet, M., Coulson, R.M.R., Cronin, A., de Villiers, E., Fraser, A., Fosker, N., Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F., Larke, N., Lord, A., Maser, P., McKellar, S., Mooney, P., Morton, F., Nene, V., O'Neil, S., Price, C., Quail, M.A., Rabbinowitsch, E., Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T., Squares, R., Squares, S., Tivey, A., Walker, A.W., Woodward, J., Dobbelaere, D.A.E., Langsley, G., Rajandream, M.A., McKeever, D., Shiels, B., Tait, A., Barrell, B., Hall, N., 2005. The genome of the host-cell transforming parasite *Theileria annulata* and a comparison with *T. parva*. Science 309, 131-133.

Petrigh, R., Ruybal, P., Thompson, C., Neumann, R., Moretta, R., Wilkowsky, S., Draghi, G., Echaide, I., de Echaide, S.T., Farber, M., 2008. Improved Molecular Tools for Detection of *Babesia bigemina*. Ann. N. Y. Acad. Sci. 1149, 155-157.

Pipano, E., Markovics, A., Kriegel, Y., Frank, M., Fish, L., 1987. Use of long-acting oxytetracycline in the immunisation of cattle against *Babesia bovis* and *B. bigemina*. Res. Vet. Sci. 43, 64–66.

Pipano, E., Shkap, V., 2000. Vaccination against Tropical Theileriosis. Ann. N. Y. Acad. Sci. 916, 484-500.

Potgieter, F.T., Els, H.J., 1977. Light and electron microscopic observations on the development of *Babesia bigemina* in larvae, nymphae and non-replete females of *Boophilus decoloratus*. Onderstepoort J. Vet. Res. 44, 213–231.

Puente, X.S., Sánchez, L.M., Overall, C.M., López-Otín, C., 2003. Human and mouse proteases: a comparative genomic approach. Nat Rev Genet. 4, 544-558.

Quesada, V., Ordóñez, G.R., Sánchez, L.M., Puente, X.S., López-Otín, C., 2009. The Degradome database: mammalian proteases and diseases of proteolysis. Nucleic Acids Res. 37(Database issue), D239-243.

Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J., Barrett, A.J., 2008. MEROPS: the peptidase database. Nucleic Acids Res. 36, D320-D325.

Riek, R.F., 1966. The life cycle of *Babesia argentina* (lignieres, 1903) (Sporozoa: Piroplasmidea) in the tick vector *Boophilus microplus* (Canestrini). Aust. J. Agric. Res. 17, 247-254.

Robinson, P.M., 1982. Theileriosis annulata and its transmission - a review. Trop. Anim. Health Prod. 141, 3-12.

Rodriguez, R.I., Trees, A.J., 1996. In vitro responsiveness of *Babesia bovis* to imidocarb dipropionate and the selection of a drug-adapted line. Vet. Parasitol. 62, 35-41.

Roos, D.S., 2005. Themes and Variations in Apicomplexan Parasite Biology. Science 309, 72-73.

Rosenthal, P.J., 2004. Cysteine proteases of malaria parasites. Int. J. Parasitol. 34, 1489-1499.

Saitou, N., Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.

Sajid, M., McKerrow, J.H., 2002. Cysteine proteases of parasitic organisms. Mol. Biochem. Parasitol. 120, 1-21.

Sajid, M., McKerrow, J.H., 2002. Cysteine proteases of parasitic organisms. Mol.

Sako, Y., Sugimoto, C., Onuma, M., 1999. Cloning of a cysteine proteinase gene of *Theileria sergenti*. J. Vet. Med. Sci. 61, 271-273.

Salih, D.A., El Hussein, A.M., Seitzer, U., Ahmed, J.S., 2007. Epidemiological studies on tick-borne diseases of cattle in Central Equatoria State, Southern Sudan. Parasitol. Res. 101, 1035-1044.

Schmidt-Christensen, A., Sturm, A., Horstmann, S., Heussler, V.T., 2008. Expression and processing of *Plasmodium berghei* SERA3 during liver stages. Cell. Microbiol. 10, 1723-1734.

Schouls, L.M., van de Pol, I., Rijpkema, S.G., Schot, C.S., 1999. Detection and identification of *Ehrlichia*, *Borrelia burgdorferi* Sensu Lato, and *Bartonella* species in Dutch *Ixodes ricinus* ticks. J. Clin. Microbiol. 37, 2215-2222.

Shkap, V., Leibovitz, B., Krigel, Y., Hammerschlag, J., Marcovics, A., Fish, L., Molad, T., Savitsky, I., Mazuz, M., 2005. Vaccination of older *Bos taurus* bulls against bovine babesiosis. Vet. Parasitol. 129, 235-42.

Shkap, V., Pipano, E., 2001. Culture-derived parasites in vaccination of cattle against tick-borne diseases. In: House, J., Kocan, K., Gibbs, P. (Eds.), Tropical Veterinary Diseases: Control and Prevention in the Context of the New World Order. Ann. N. Y. Acad. Sci. 154–171.

Sibley, L.D., 2004. Intracellular Parasite Invasion Strategies. Science 304, 248-253.

Sijwali, P.S., Brinen, L.S., Rosenthal, P.J., 2001a. Systematic optimization of expression and refolding of the *Plasmodium falciparum* cysteine protease falcipain-2. Protein Expr. Purif. 22, 128-134.

Sijwali, P.S., Shenai, B.R., Gut, J., Singh, A., Rosenthal, P.J., 2001b. Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3. Biochem J. 360, 481-489.

Singh, A., Shenai, B.R., Choe, Y., Gut, J., Sijwali, P.S., Craik, C.S., Rosenthal, P.J., 2002. Critical role of amino acid 23 in mediating activity and specificity of vinckepain-2, a papain-family cysteine protease of rodent malaria parasites. Biochem. J. 368, 273-281.

Smith HV, Nichols RA, Grimason AM., 2005. *Cryptosporidium* excystation and invasion: getting to the guts of the matter. Trends Parasitol. 21, 133-142.

Smith, T., Kilborne, F.L., 1893. Investigations into the nature, causation and prevention of Southern cattle fever. In Ninth Annual Report of the Bureau of Animal Industry for the Year 1892, pp. 177–304. Washington, Government Printing Office.

Sneath, P.H.A., Sokal, R.R., 1973. Numerical Taxonomy. Freeman, San Francisco.

Spooner, R.L., Innes, E.A., Glass, E.J., Brown, C.G., 1989. *Theileria annulata* and *T. parva* infect and transform different bovine mononuclear cells. Immunology 66, 284-288.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596-1599.

Taylor, R.J., Mchardy, N., 1979. Preliminary observations on the combined use of imidocarb and *Babesia* blood vaccine in cattle. J. S. Afr. Vet. Assoc. 50, 326–329.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

Trees, A.J., 1974. The application of acridine orange staining to quantitate low levels of *Babesia divergens* parasitaemias. Trans. R. Soc. Trop. Med. Hyg. 68, 277.

Uilenberg, G., 2006. *Babesia* - a historical overview. Vet. Parasitol. 138, 3-10.

Van Niekerk, C.J., Zweygarth, E., 1996. *In vitro* cultivation of *Babesia occultans*. Onderstepoort J. Vet. Res. 63, 259-61.

Vial, H.J., Gorenflot, A., 2006. Chemotherapy against babesiosis. Vet. Parasitol. 138, 147-160.

Wallace, R.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., Itakura, K., 1979. Hybridization of synthetic oligodeoxyribonucleotides to PhiX 174 DNA: the effect of single base pair mismatch. Nucl. Acids Res. 6, 3543-3557.

Wickham, M.E., Culvenor, J.G., Cowman, A.F., 2003. Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. J. Biol. Chem. 278, 37658-37663.

Wilkie, G.M., Brown, C.G., Kirvar, B.E., Thomas, M., Williamson, S.M., Bell-Sakyi, L.J., Sparagano, O., 1998. Chemoprophylaxis of *Theileria annulata* and *Theileria parva* infections of calves with buparvaquone. Vet. Parasitol. 78, 1-12.

Winn Jr., W.C., Allen, S.D., Janda, W.M., Koneman, E.W., Procop, G., Schreckenberger, P.C., Woods, G.L., 2006. Color Atlas and Textbook of Diagnostic Microbiology. 6th edition, Lippincott Williams & Wilkins.

Wright, I.G., 1990. Immunodiagnosis and immunoprophylaxis against the haemoparasite *Babesia* sp. and *Anaplasma* sp. in domestic animals. Rev. Sci. Tech. Off. Int. Epiz. 9, 345-356.

Wu, Y., Wang, X., Liu, X., Wang, Y., 2003. Data-Mining Approaches Reveal Hidden Families of Proteases in the Genome of Malaria Parasite. Genome Res. 13, 601-616.

Yeoh, S., O'Donnell, R.A., Koussis, K., Dluzewski, A.R., Ansell, K.H., Osborne, S.A., Hackett, F., Withers-Martinez, C., Mitchell, G.H., Bannister, L.H., Bryans, J.S., Kettleborough, C.A., Blackman, M.J., 2007. Subcellular Discharge of a Serine Protease Mediates Release of Invasive Malaria Parasites from Host Erythrocytes. Cell 131, 1072-1083.

Yin, H., Luo, J., Lu, W., 2008. Control of tropical theileriosis with attenuated schizont vaccine in China. Vaccine 26S, G11-G13.

Zintl, A., Mulcahy, G., Skerrett, H.E., Taylor, S.M., Gray, J.S., 2003. *Babesia divergens*, a bovine blood parasite of veterinary and zoonotic importance. Clin. Microbiol. Rev. 16, 622-636.

Thesis Publications

Martins, T.M., Pedro, O.C., Caldeira, R.A., do Rosário, V.E., Neves L., Domingos, A., 2008. Detection of bovine Babesiosis in Mozambique by a novel seminested hot-start PCR method. Veterinary Parasitology 153, 225–230.

Martins, T.M., Neves L., Pedro, O.C., Fafetine, J.M., do Rosário, V.E., Domingos, A., 2009. Molecular study on bovine babesiosis and other cattle haemoparasites infections in Maputo province, Mozambique. (submitted)

Martins, T.M., do Rosário, V.E., Domingos, A., 2009. Identification of papain-like cysteine proteases from the bovine piroplasm *Babesia bigemina*. (submitted)

Martins, T.M., do Rosário, V.E., Domingos, A., 2009. Expression and characterization of the *Babesia bigemina* cysteine protease babesipain-1. (submitted)

Martins, T.M., Gonçalves, L.M.D., Capela, R., Moreira, R., do Rosário, V.E., Domingos, A., 2009. Effect of synthesised inhibitors on babesipain-1, a new cysteine protease from the bovine piroplasm *Babesia bigemina*. (submitted)